

**UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
INSTITUTO DE INGENIERÍA
MAESTRÍA Y DOCTORADO EN CIENCIAS E INGENIERÍA**



**“IMPLEMENTACIÓN DE TÉCNICAS DE SOCIAL MEDIA PARA EL
ANÁLISIS DE DATOS EN FACEBOOK”**

**TESIS QUE PARA OBTENER EL GRADO DE
DOCTOR EN CIENCIAS**

PRESENTA:

Paola Elvira Velázquez Solís

DIRECTORA:

Dra. Brenda Leticia Flores Ríos

CODIRECTORA:

Dra. Mónica Carrillo Beltrán

DEDICATORIA

A esas personas especiales que siempre creyeron que podía hacer esto,
pero ahora ya no están aquí conmigo para celebrarlo.

Para mi tío Jesús, mi Mamá Elvira
y muy especialmente para ti, Abuelo.

AGRADECIMIENTOS

Primeramente, a Dios.

A mi familia, principalmente a mis padres, mi hermano, mi prometido y todos aquellos que se acercaron a mí con una palabra de aliento cuando creía que no podía más.

A mis amigos y compañeros de trabajo por esas reuniones extra oficiales donde eran capaces de darme claridad en mis ideas y aportar a este proyecto.

A los miembros del sínodo, Dra. Mónica Carrillo Beltrán, Dr. Jorge Eduardo Ibarra Esquer, Dra. María Angélica Astorga Vargas, Dr. Jesús Caro Gutiérrez y Dra. Yamel Ungson, por aceptar revisar y brindar comentarios valiosos y por su contribución general a este trabajo. Así mismo, al Dr. Raúl Antonio Aguilar Vera de la Universidad Autónoma de Yucatán por todo el apoyo como revisor externo y la colaboración en estos últimos tres años.

Al Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT), por la beca otorgada para la realización de mis estudios de doctorado.

Finalmente, quisiera agradecer a mi directora, Dra. Brenda Leticia Flores Ríos, por no sólo ser una guía desde que iniciamos con este proyecto, sino ir más allá y hacerme sentir parte de una familia académica.

RESUMEN de la tesis de **PAOLA ELVIRA VELAZQUEZ SOLIS**, presentada como requisito, presentada como requisito para la obtención del **GRADO DE DOCTOR EN CIENCIAS**. Mexicali, Baja California, México. Enero de 2024.

IMPLEMENTACIÓN DE TÉCNICAS DE SOCIAL MEDIA PARA EL ANÁLISIS DE DATOS EN FACEBOOK

Se presenta **IMPLEMENTACIÓN DE TÉCNICAS DE SOCIAL MEDIA PARA EL ANÁLISIS DE DATOS EN FACEBOOK**, como una propuesta a las organizaciones que se encuentran en la búsqueda de mejorar la participación de los usuarios en las redes sociales, con el objetivo de promover la visibilidad de sus propias actividades. Es así como a través de esta presencia en internet, las actividades y sus resultados se convierten en activos valiosos más allá de los entornos de negocio. Como resultado, la relación con la comunidad se fortalece, abriendo nuevas oportunidades de colaboración y permitiendo adaptar el contenido con la interpretación de las tendencias. Los datos obtenidos de la creación de contenido y la interacción en las redes sociales permiten identificar los elementos de User Engagement, interactividad, visibilidad y alcance. A partir de estos elementos, es posible implementar técnicas de Social media para comprender el comportamiento de los usuarios y mejorar el alcance del contenido. Se realizó un análisis estadístico, cuantitativo y cualitativo de métricas de Facebook, como las reacciones, los clicks y variables como el User Engagement desde una perspectiva de compromiso universidad-comunidad. Se siguió una metodología de intervención para identificar cómo cada elemento contribuye a la participación del usuario y se ha obtenido como resultado, el Marco de trabajo para el minado de Social media basado en GC, como una herramienta estratégica alternativa para mejorar las prácticas de creación y difusión de contenidos en torno a métricas como el User Engagement. El estudio se realizó a partir de tres muestras de datos de fan pages de Facebook utilizadas para la difusión de contenidos, noticias y eventos científicos y culturales. Las variaciones en las métricas de las publicaciones individuales se explican mediante dos modelos de regresión, creados con las variables más representativas extraídas del informe de interacción de la fan page y con un análisis más detallado; utilizando técnicas de minería de datos se muestra que el contenido de producción propia es el factor principal para la mayor participación de la comunidad. El marco de trabajo tiene la capacidad de permitir que los roles que participa en crear y administrar los contenidos de redes sociales comprendan el alcance y los resultados tanto de los contenidos como

de los usuarios de su comunidad. Se puede utilizar tanto para interpretar los datos existentes como para diseñar estrategias de seguimiento, que es donde los roles asociados a crear y administrar contenidos se beneficiarían.

TÉRMINOS CLAVE: Social media, Engagement, Interactividad, Modelos de regresión.

ÍNDICE

DEDICATORIA	2
AGRADECIMIENTOS	3
ÍNDICE	6
ÍNDICE DE FIGURAS.....	8
ÍNDICE DE TABLAS	8
CAPÍTULO I: INTRODUCCIÓN	1
1.1 Definición del problema.....	3
1.2 Justificación.....	4
1.3 Preguntas de Investigación.....	8
1.4 Objetivos	9
1.4.1 Objetivo General.....	9
1.4.2 Objetivos Específicos	9
1.5 Metodología de investigación	9
1.6 Estructura del documento.....	10
CAPÍTULO II: MARCO REFERENCIAL	12
2.1 Redes sociales	12
2.2 Social media	13
2.2.1 Características.....	14
2.2.2 Técnicas	15
2.2.3 Procesos	18
2.3 Algoritmo en un enfoque de minería de datos	19
2.5 User Engagement en Facebook.....	21
2.5.1 Características.....	21
2.6 Identificación de las áreas de aplicación de Social media apoyadas en modelos, técnicas y herramientas	23
2.6.1 Identificación de modelos para Social media en diversas perspectivas	24
2.6.2 Técnicas aplicadas y herramientas implementadas en Social media.....	25
2.6.3 Análisis del conjunto de datos	26
CAPÍTULO III: MARCO DE TRABAJO PARA EL MINADO DE SOCIAL MEDIA BASADO EN GC	28
3.1 Metodología de intervención.....	28

3.1.1 Auditoría de conocimiento	28
3.1.2 Identificación de herramientas tecnológicas.....	30
3.1.3 Despliegue del sistema de GC	31
3.1.4 Seguimiento y evaluación.....	32
3.2 Descripción del Marco de trabajo para el minado de Social media basado en GC.....	33
3.2.1 Proceso 1: Preparación y preprocesamiento de los datos	35
3.2.2 Proceso 2: Implementación de técnicas y métodos	38
3.2.3 Proceso 3: Análisis de los datos de redes sociales.....	40
3.2.4 Proceso 4: Validación por estudio de caso	41
CAPÍTULO IV: EXPERIMENTACIÓN	44
4.1 Fan page Mujeres en la Investigación - UABC (Parte 1).....	44
4.1.1 Análisis de las variables	44
4.1.2 Modelo de regresión y Minería de datos	50
4.2 Fan page Mujeres en la Investigación - UABC (Parte 2).....	53
4.2.1 Análisis de las variables	54
4.2.2 Modelo de regresión y Minería de datos	57
4.3 Cultura UABC.....	59
4.3.1 Análisis de las variables	60
4.3.2 Análisis de la comunidad.....	61
CAPÍTULO V: CONCLUSIONES, APORTACIONES Y TRABAJO FUTURO	66
5.1 Conclusiones	66
5.2 Aportaciones.....	68
5.2 Trabajo futuro.....	68
REFERENCIAS.....	70
Anexo 1. Lista de productos académicos.....	78
Anexo 2. Participación en eventos académicos	81

ÍNDICE DE FIGURAS

Figura 1 Modelo de investigación con enfoque transdisciplinario.	10
Figura 2 Procesos de Social media. Adaptado de: Goyal & Goyal (2020); Meng et al. (2020); Mushtaq et al. (2018); Gupta & Gupta (2016); Liu et al. (2016).....	19
Figura 3 Proceso de Selección de artículos.....	23
Figura 4 Gráfico de dispersión de puntos con la descripción de los estudios primarios.	27
Figura 5 Identificación de roles de redes sociales y tipos de conocimiento.	29
Figura 6 Modelo global para el diseño de un marco de trabajo basado en Social media.	33
Figura 7 Flujo de trabajo de las actividades de los cuatro procesos de Social media Mining Framework.	35
Figura 8 Marco de trabajo para el minado de Social media basado en GC.	43
Figura 9 Diagrama de flujo de Engagement para fan page en Facebook.	46
Figura 10 Diagrama de dispersión de las variables analizadas.	48
Figura 11 QQ-Graphs para verificación de normalidad de datos.	49
Figura 12 Gráfico de burbujas con los tipos de contenidos que comparte la fan page.	52
Figura 13 Gráfica de variables extraídas de Facebook con la técnica de k-medias para k=3.....	53
Figura 14 Distribución por edades de los seguidores de las páginas de fans.	54
Figura 15 User Engagement con las dos fórmulas.....	58
Figura 16 Agrupamiento de User Engagement con las dos fórmulas.....	58
Figura 17 Histograma de Engagement.....	61
Figura 18 Metodología para la identificación de comunidades en Facebook.....	62
Figura 19 Grafos generados por Gephi.....	63
Figura 20 Cloud Word de la comunidad en Facebook para la Cultura y Divulgación de la Ciencia.	65

ÍNDICE DE TABLAS

Tabla 1 Algoritmos de distribución seleccionados en Gephi (Bedi & Sharma, 2016; Diliانا & Indrawati, 2022; Humberstone, 2019).	17
Tabla 2 Definiciones de algoritmo para diversas áreas de aplicación.	20
Tabla 3 Fórmulas para calcular el Engagement en Facebook.....	22
Tabla 4 Estudios con áreas de identificación con orientación a Social media.....	24

Tabla 5 Estudios con Modelos de análisis basados en Social media.	25
Tabla 6 Estudios enfocados en Arquitecturas de software basadas en Social media.	25
Tabla 7 Estudios de técnicas y herramientas basadas en Social media.	26
Tabla 8 Asociación de tecnologías con Procesos de conocimiento desde la perspectiva de Social media.	30
Tabla 9 Ámbitos de actuación de los Social media en las áreas estratégicas.	31
Tabla 10 Especificación de los tres procesos del Social media Mining Framework.	34
Tabla 11 Casos atípicos en el preprocesado de los datos.	37
Tabla 12 Métodos, técnicas y herramientas sugeridas por cantidad de tuplas.	39
Tabla 13 Ejemplo de datos obtenidos desde Facebook.	45
Tabla 14 Diagrama de flujo de Engagement para fan page en Facebook.	47
Tabla 15 Variables de categorización de contenidos.	50
Tabla 16 Tabla de contingencia para la prueba de normalidad chi-cuadrado.	50
Tabla 17 Distribución de reacciones, comentarios y acciones extraídas del conjunto de datos. ..	55
Tabla 18 Porcentaje de publicaciones en las categorías.	56
Tabla 19 Comparación de los dos años entre las fórmulas por mes.	57
Tabla 20 Comparación de correlaciones de Spearman por categoría.	57
Tabla 21 Descripción de muestra del programa de extensión de la cultura y la Ciencia.	60
Tabla 22 Estadísticos básicos.	60

CAPÍTULO I: INTRODUCCIÓN

Las redes sociales son utilizadas por más de 2.95 billones de personas (Liyanage & Premarathne, 2021) que interactúan entre sí en una plataforma digital y han cambiado la forma en que se crea contenido, comparte y comunica información y los usuarios se relacionan (Froment et al., 2017). Las plataformas de redes sociales se destacan por contar con una extensa y diversa base de usuarios, con participación constante e intermitente, convirtiéndose en un medio esencial para la comunicación y la obtención de información (Ureña et al., 2019; Velazquez-Solis et al., 2021). En contraste, Social media y hace referencia al conjunto de métodos, herramientas, aplicaciones y medios de comunicación donde los usuarios crean y visualizan información, teniendo la libertad de compartir y transferir diversos tipos de contenido, como textos, fotografías, audio y vídeo (Injadat et al., 2016). La integración de criterios de análisis y técnicas características de Social media en las plataformas de redes sociales tiene como objetivo facilitar una comunicación ubicua mediante la creación y compartición de contenidos (Martín Blanco, 2015).

Los datos producidos por las redes sociales son una base de evidencia de gran volumen, variada en contenido y dinámicas en el comportamiento de usuarios, proporcionando nuevas oportunidades para comprender a las personas, los grupos escolares, académicos, deportivos y la sociedad misma. Por ejemplo, han adquirido una importancia fundamental para ciertas organizaciones, especialmente en áreas como el servicio al cliente o los departamentos de marketing, donde se desarrollan soluciones para la identificación de influencers¹ y los nichos de disponibles en el ámbito digital (Arora et al., 2019). Estas organizaciones proponen estrategias efectivas de publicación en diversas redes sociales, con el objetivo de reconocer el tipo de mensajes y los métodos de transmisión de mensajes adecuados (Bitiktas & Tuna, 2020).

Se ha detectado un creciente interés entre científicos y profesionales por compartir sus investigaciones utilizando diversas redes sociales (Amin & Rajadurai, 2018; Ovadia, 2014). Algunos de los sitios de redes sociales, también llamadas redes profesionales (Bitiktas & Tuna, 2020; Herrera, 2012; Celaya, 2008) con enfoque de investigación y académico son: ResearchGate, Academia.edu y Mendeley, las cuales son redes sociales que los académicos e investigadores usan

¹ Persona(s) que sobresale(n) en una plataforma de redes sociales u otro medio de comunicación, emitiendo puntos de vista sobre un tema específico que ejercen una gran influencia sobre un considerable número de personas familiarizadas con su presencia.

para compartir resultados parciales o completos de sus investigaciones (Singson & Amees, 2017). Sin embargo, los estudiantes son los usuarios más frecuentes y representan el segmento más grande en el uso de Internet (Olea Méndez & Moscosa Ruiz, 2018; Yaakop et al., 2013). Por lo tanto, los académicos están aprovechando los sitios de redes sociales diseñados con fines más orientados a la comunicación y negocios, como Facebook o Twitter, para mejorar su visibilidad, conectarse y compartir el conocimiento que generan (Al-Maatouk et al., 2020; Singson & Amees, 2017).

Los científicos y profesionales de las organizaciones descubren a diario nuevas formas de recopilar la información contenida en los datos, combinar los diferentes tipos de conocimiento y de forma automática analizar los datos desde un enfoque organizacional (Batrinca & Treleaven, 2015). Los modelos actuales se enfocan en extraer conocimiento de conjuntos de datos para analizar no solo comportamientos, sino también emociones y sentimientos (He et al., 2017; Salloum et al., 2017). Por tanto, es esencial proporcionar un marco de trabajo que permita aprovechar los datos no estructurados e interpretarlos, facilitando su exploración y explotación adecuadas (Injadat et al., 2016).

En este contexto, algunos expertos en Ciencia de datos desarrollan algoritmos fundamentados en los modelos propuestos, con el objetivo de continuar supervisando el rendimiento de la información creada en las redes sociales, así minimizando el tiempo necesario para conocer las métricas completas de páginas en Facebook o Twitter (Purohit et al., 2020; Kurnia & Suharjito, 2018). En este contexto, las redes sociales, los Social media y los entornos virtuales desempeñan funciones como canales de comunicación y mecanismos de interacción, con un potencial de impacto y alcance que superan a los medios tradicionales, y un dinamismo y capacidad de interacción que los sitúa en un nivel superior (Herrera, 2012; Velazquez-Solis et al., 2021). Así mismo, son una realidad de conjunto de cambios, tanto a nivel individual como institucional. Por consiguiente, implementar técnicas de Social media para el análisis de las redes sociales, como parte de herramientas para resolver problemas comunes en la interpretación de datos, proporcionará un marco de referencia para guiar un análisis enfocado en necesidades de los usuarios.

1.1 Definición del problema

Los Social media se han convertido en un grupo importante de datos no estructurados que pertenecen a una gran cantidad de dominios, los cuales incluyen áreas como la educación, los negocios, el gobierno y la salud. Sin embargo, el contenido que se genera a partir de los usuarios en las redes sociales permanece en gran medida sin explorar, ni explotar (Velazquez-Solis, 2017; Asur & Huberman, 2010). Por ello, el uso de herramientas implementadas en Social media está aumentando masivamente en las organizaciones, pero aún carece de propósito de adopción. Por ejemplo, las herramientas de redes sociales no se utilizan con fines educativos, siendo más orientadas a las demandas o necesidades de los usuarios (Singh et al., 2022; Adinolfi et al., 2016). El problema real es la naturaleza de las redes sociales y las políticas institucionales, que crean un conflicto que a su vez provoca una brecha de transferencia de conocimiento (Amin & Rajadurai, 2018). Las complicaciones de las redes sociales en relación con el aprendizaje y la educación no son nuevas y se han investigado a fondo para instituciones que tienen sus propios valores culturales, políticas y jerarquías a las que están vinculadas. Por ello, es necesario entender que, con el cambio de tecnología, las instituciones académicas deben adoptar estas herramientas, pero con sus propios términos (Amin & Rajadurai, 2018; Rodriguez-Vazquez et al., 2016). Además, es crucial reconocer la necesidad de desarrollar estrategias efectivas de publicación en las redes sociales, que incluyan métodos para identificar el tipo de mensajes y los métodos adecuados de transmisión de información (Bitiktas & Tuna, 2020) con el objetivo de facilitar a docentes y académicos de diversas disciplinas y culturas la capacidad de impartir cursos y compartir información, se recurre a plataformas como las redes sociales. Esto se logra gracias a las variadas funcionalidades proporcionadas por herramientas de Social media, tales como las historias de voz en Twitter, el Portal Home y la nube de Facebook. Estas herramientas han motivado a los estudiantes a participar de manera más activa en sus actividades académicas (Rousidis et al., 2020).

La creciente dependencia de las redes sociales exige técnicas de extracción de datos que faciliten la interpretación y aprovechamiento de los datos no estructurados y los ubiquen dentro de un patrón sistemático en las bases de datos y/o bases de conocimiento (Injadat et al., 2016). Las organizaciones utilizan como repositorios las bases de datos, bases de conocimiento, intranet o sistemas de información, donde el tamaño y uso de estas se basa en criterios sobre capacidad de almacenamiento y eficiencia de procesamiento de información, dejando de lado el uso posterior

de la información contenida para la exploración, explotación y análisis (Velazquez-Solis, 2017). Se ha detectado la necesidad de implementar la información contenida en estos repositorios en el análisis del comportamiento social virtual, como parte del conjunto de técnicas que proporcionen un marco de referencia para guiar el uso de los datos de las redes sociales que no solo apoye en la vinculación, divulgación y promoción de la transferencia de diversos tipos de conocimientos generados por una organización educativa, sino también, aproveche la información generada diariamente. El diseño y la implementación de este marco de trabajo lograrán satisfacer necesidades informacionales en relación con la actividad científica. El apropiado análisis de los datos destinados a difundir el conocimiento de los centros de Educación Superior contribuiría a la formación en investigación científica (Hernández-García et al., 2020).

1.2 Justificación

La conjunción entre el auge de la educación en línea y la creciente utilización de dispositivos conectados a internet por parte de los estudiantes ha contribuido a enriquecer los entornos de aprendizaje. Esto ha propiciado la interactividad en entornos cada vez más adaptados a las necesidades individuales del usuario (Prada Núñez et al., 2019). La interactividad es una representación de las preferencias del usuario y la forma en que perciben la información, alimentando algoritmos con datos para futuras recomendaciones. Se asume que un alto nivel de interactividad genera diversos tipos de actitudes sobre los contenidos de las fan pages o grupos que existen en Facebook. Estas actitudes pueden ser positivas o negativas (Kaye, 2021). Todo ello limitado por la comparación entre los resultados de la interactividad en diferentes publicaciones de la misma naturaleza (Ballesteros Herencia, 2018; Sutanto, 2016).

Las redes sociales permiten a través de internet que los usuarios puedan (Peruta & Shields, 2017): 1) establecer un perfil semipúblico o público, 2) crear una lista de otros usuarios con los que comparten conexiones y contenido, y 3) explorar un conjunto de datos vinculados a las conexiones y contenidos generados por otros usuarios en el sistema. De este modo, una red social se convierte en una herramienta que cambia el rol de las personas, haciéndolas tanto consumidores como productores de contenido (Gonzalez & Munoz, 2016; Velazquez-Solis et al., 2021), generando nuevas formas de interacción y configurando innovadores métodos de comunicación

entre individuos. Esto abarca la toma de decisiones, la socialización, la colaboración, el aprendizaje, las transacciones comerciales y la interacción entre usuarios (Sabate et al., 2014). Las áreas estratégicas cruciales a monitorear en los Social media incluyen las necesidades del usuario, las emociones o sentimientos y la comunicación entre ellos. Aunque este seguimiento se realiza inmediatamente en Facebook (Kurnia & Suharjito, 2018), se reconoce como imperante la creación de marcos de trabajo que ofrezcan una metodología centrada en los Social media (Purba et al., 2020; Verhoef & Lemon, 2013).

La cultura científica busca crear trazabilidad con los objetivos de las organizaciones, válidos para todos los públicos, vinculados al presente y futuro de la Ciencia y la Tecnología. De esta forma, la información científica, especialmente si se produce desde las Instituciones de Educación Superior (IES) debe aunar estos aspectos y ofrecer cultura científica a la población, para generar conocimiento (Alsufyan & Aloud, 2017; Toharia, 2010). Por lo tanto, la información científica debe ser considerada como un instrumento importante para el desarrollo de las nuevas Redes sociales y la difusión institucional del conocimiento (Campos-Freire & Rúas-Araújo, 2016). El uso más común de las Redes sociales en la información científica es compartir el conocimiento de las publicaciones de otros investigadores, dar a conocer las propias y seguir a otros investigadores con líneas de investigación afines. De esta forma, se ha modificado la dinámica de la comunicación científica, acelerando el ciclo de retroalimentación del conocimiento (Orioque, 2021; Peruta & Shields, 2017; Campos-Freire & Rúas-Araújo, 2016) gracias a las redes sociales. Por ello, se considera que el análisis de grandes volúmenes de datos con redes sociales es una tendencia emergente que tiene posibilidad de emplearse en numerosas aplicaciones de la vida real. Recientemente, los investigadores y las comunidades académicas se ocupan de cuestiones y estudios en evolución de las ciencias sociales y de la informática (Rousidis et al., 2020). Esta popularidad y adopción por parte de las masas ha cambiado la forma en que se llevan a cabo los análisis de los grandes volúmenes de datos. Son una herramienta que impactaría en un cambio positivo en el sector educativo a la forma en que proporciona información y ayuda a conectar personas e ideas, actuando así como un agente que recopila las contribuciones de personas de todo el mundo (Prestridge et al., 2021; Dabbagh & Kitsantas, 2012).

Los análisis de las principales diferencias entre instituciones públicas y privadas en las redes sociales, en concreto Facebook, arrojan resultados positivos sobre la incorporación de

elementos de este tipo en la profesionalización e institucionalización de las organizaciones (Rodríguez-Vazquez et al., 2016). Existen algunos reportes de casos de estudio en universidades, como el de González & Muñoz (2016), en el que se identificó principalmente a la población estudiantil y qué tipo de necesidades e interacción tenían con las redes sociales; para posteriormente buscar estrategias de enseñanza-aprendizaje incorporando las redes sociales como herramienta. Numerosos académicos están recurriendo a las redes sociales para comunicarse con los estudiantes, ya que la credibilidad parece aumentar en los estudiantes, resultando en que estén más comprometidos con la publicación de la actividad académica del maestro en las redes sociales (Rousidis et al., 2020). En el estudio comparativo de Brito et al. (2012) participaron cinco universidades iberoamericanas de Argentina (Universidad Nacional de Buenos Aires y la Universidad Nacional de Quilmes), México (Universidad de Guadalajara y el Instituto Tecnológico de Monterrey) y España (Universidad Nacional de Educación a Distancia), con el propósito de proponer una estrategia concisa del uso de las redes sociales a nivel institucional bajo un esquema de relacionar las actividades en las redes sociales con los objetivos de la universidad, en la docencia, la investigación, la administración; y específicamente, mostrar cómo aumentar su: capital intelectual, capital relacional y capital de identidad. Al tratar de aumentar el alcance de sus publicaciones, las IES varían con frecuencia el contenido de sus publicaciones, con la esperanza de aumentar el compromiso de sus followers² o seguidores. Para determinar si existe una combinación de contenido ideal, aunque es bien sabido que primero es importante comprender el panorama actual (Peruta & Shields, 2017). Aunque, las Redes sociales se han convertido en una implementación en plataformas de cambio social (Huertas et al., 2015) están adquiriendo cada vez más importancia como canal para recopilar información sobre productos, servicios y nuevas oportunidades (Penttinen & Frösén, 2022; Verhoef & Lemon, 2013). Por ello, las universidades tienen que aprovechar el potencial que tienen en todos los aspectos (Sánchez-Rodríguez et al., 2015). Por lo tanto, una IES debe afrontar con más fuerza su responsabilidad social, ya que sus redes sociales se convierten en un espacio de conocimiento e investigación global, con docentes, estudiantes y profesionales conocedores del uso educativo de las tecnologías de la información y la comunicación con las redes sociales (Buxarrais, 2016).

² Usuario que se ha suscrito al perfil de otros usuarios, en particular en el servicio de micro-blogging Twitter o la red social Facebook.

Un enfoque cualitativo para el análisis de datos de redes sociales apoya en gran medida la creación de significado al colocar la tecnología en contextos sociales, lugares y tiempos específicos. Por otro lado, existe la necesidad de implementar métodos inductivos que permitan profundizar en la cultura de los espacios digitales, considerando las perspectivas de los usuarios como complejas y heterogéneas (Wheatley & Vatnoey, 2020). Una revisión de trabajos relacionados identificó la aplicación de varios tipos de métodos, que establecieron datos empíricamente usando métodos estadísticos aceptados (Karademir Coşkun et al., 2020; Voorveld et al., 2018; Natthaphong Phuntusil & Yachai Limpiyakorn, 2017), como pruebas no paramétricas (Molina et al., 2021; He et al., 2017), análisis ANOVA de varianza y análisis de correlación de coeficientes (Gil-quintana et al., 2022; Velazquez-Solis et al., 2022), modelos de análisis estadístico (Voorveld et al., 2018), modelos de regresión (Wheatley & Vatnoey, 2020), modelos de asociación, modelos predictivos o de minería de datos como Naive Bayes (He et al., 2017), Análisis de componentes principales (Gil-Quintana et al., 2022) y K-means (Liyange & Premarathne, 2021). Los coeficientes de correlación de Pearson, la prueba ANOVA, la prueba de chi-cuadrado y el discriminante lineal se utilizaron para un análisis del conjunto de datos (Agusriandi et al., 2020). Los modelos propuestos están basados en extraer el conocimiento de los conjuntos de datos por medio de minería de texto o métodos de estadística, para analizar los datos posteriormente con el fin de comprender e interpretar el comportamiento social virtual, así como los emociones y sentimientos de las personas que interactúan (Tsvetkov, 2021; Wheatley & Vatnoey, 2020). Existen propuestas metodológicas que presentan aplicaciones de algoritmos para análisis de contenido de datos de texto (Chan-Olmsted et al., 2017), algoritmo PageRank y algoritmo HITS, API de Facebook (Tsinovoi, 2020), JavaScript y API REST (Dabbagh & Kitsantas, 2012). En algunos de ellos, las actividades de extracción y limpieza de datos en el proceso de análisis de datos de redes sociales son similares (Pilař et al., 2021; Surya Gunawan et al., 2020; Eriksson et al., 2019; Natthaphong; Stieglitz et al., 2018; Phuntusil & Yachai Limpiyakorn, 2017).

En el trabajo de Franz et al. (2019), se realizó una investigación cualitativa sobre los usuarios de Facebook y su actividad en la red social. Buscaron opciones para abordar temas de salud a través del análisis de los textos generados en la red social, concluyendo que el análisis cualitativo y cuantitativo en Facebook involucra la participación del investigador en el estudio de patrones de información observados para identificar temas relacionales, elementos gramaticales y

el valor de las emociones contenidas en las publicaciones de Facebook y los comentarios asociados. En consecuencia, las redes sociales, Social media y los espacios virtuales son mecanismos de interacción, canales de comunicación, con un alcance y potencial de impacto mucho mayor que cualquiera de los otros medios tradicionales, con capacidad de interacción y dinamismo, superior a cualquiera de los demás medios (Herrera, 2012). Son una realidad en todos los cambios, tanto a nivel individual como institucional. Basándonos en lo previamente mencionado, se destaca la importancia de integrar las técnicas de los medios sociales como componente esencial dentro de un marco de trabajo diseñado para extraer y analizar datos provenientes de las redes sociales. El propósito central de este estudio consiste en reconocer modelos de análisis, las técnicas disponibles y las herramientas adecuadas de los Social media empleados en redes sociales como Facebook, tal como han sido propuestos en diversas áreas de aplicación. Este objetivo se aborda mediante la ejecución de un mapeo sistemático de la literatura.

1.3 Preguntas de Investigación

Del planteamiento anterior, surge la motivación para la investigación, la cual parte de la siguiente pregunta inicial:

- **PI_1** ¿Cómo desarrollar un marco de referencia que permita implementar técnicas tomando en cuenta la información analizada por los Social media?

La cual, debe satisfacer a las siguientes sub-preguntas:

- **PI_2** ¿Qué elementos (Conceptos, procesos, modelos, técnicas y herramientas) se han implementado en los Social media para aprovechar la información obtenida por las Redes sociales (Educación, Investigación, Marketing)?
- **PI_3** ¿Qué tipo de información existe en las fan page y grupos de Facebook que me permita conocer el comportamiento social virtual y el dominio del conocimiento?
- **PI_4** ¿Qué elementos son necesarios para construir un marco de trabajo basado en las técnicas y las herramientas de Social media para el análisis social virtual?
- **PI_5** ¿Cómo implementar un marco de trabajo basado en las técnicas y las herramientas de Social media para el análisis de datos extraídos de Facebook?

1.4 Objetivos

1.4.1 Objetivo General

El objetivo general de esta investigación es crear un marco de trabajo que permita implementar técnicas de Social media para análisis de datos de extraídos de Facebook.

1.4.2 Objetivos Específicos

1. Comparar el contexto del estudio de caso y el sujeto de investigación seleccionado.
2. Realizar una revisión de literatura que permita situar la problemática dentro de las disciplinas de Inteligencia Artificial (Minería de datos), Ingeniería de Software y Gestión de conocimiento para fortalecer los conceptos relacionados.
3. Identificar el tipo de información que existe en la red social Facebook para obtener datos históricos y comportamiento social virtual.
4. Desarrollar un marco de trabajo para implementar técnicas de Social media de manera que se garantice su funcionamiento y extensibilidad dentro de cualquier estudio de caso.
5. Aplicar un algoritmo para calcular interactividad y Engagement que permita realizar actividades de difusión, enseñanza y apropiación de la Ciencia y la Tecnología, impulsando la igualdad de género (Caso programa Mujeres en la Investigación, del Instituto de Ingeniería de la UABC).
6. Realizar una validación del Marco de trabajo a los conjuntos de datos originados en diferentes periodos dentro de fan pages de Facebook.

1.5 Metodología de investigación

Para dar cumplimiento a las preguntas de investigación planteadas en la sección 1.3.1, se llevaron a cabo los siguientes pasos metodológicos:

1. Se llevó a cabo un estudio comparativo de la presencia femenina a nivel nacional de estadísticas de mujeres graduadas en posgrados de química, física y medio ambiente, como consumidores de contenido del estudio de caso.
2. Se realizó un mapeo sistemático de la literatura de las técnicas, métodos y herramientas implementadas en Social media para identificar las diferentes técnicas, métodos y herramientas asociados a las áreas de aplicación (Marketing, Gestión de negocios, Educación e Investigación).
3. Se aplicó la metodología de intervención desde modelos de gestión de conocimiento para la identificación de roles, áreas estratégicas de flujos de conocimiento, los procesos y las herramientas asociadas al análisis de datos en Social media.
4. Se definió un marco de trabajo enfocado en el minado de los datos de Social media.
5. Se validó por medio del marco de trabajo un estudio de caso para analizar los datos de una fan page de Facebook con enfoque a la difusión de información científica.

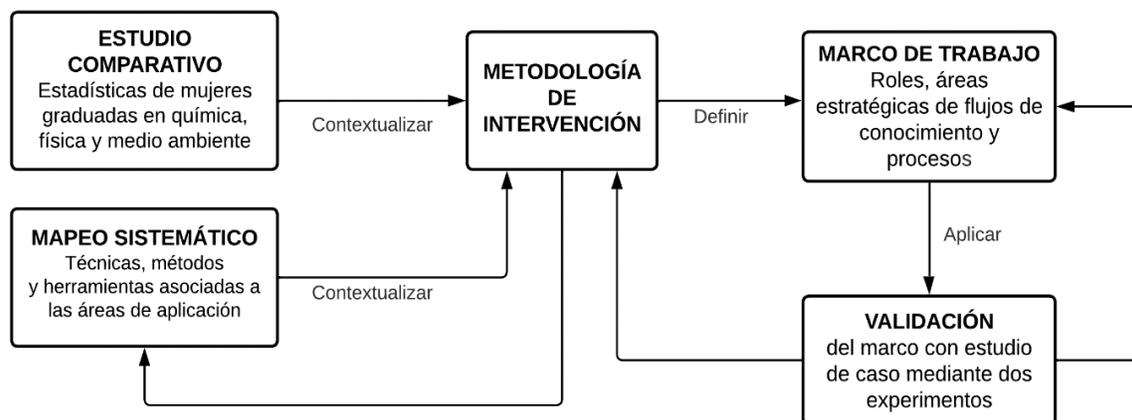


Figura 1 Modelo de investigación con enfoque transdisciplinario.

1.6 Estructura del documento

El resto del documento está organizado de la siguiente manera:

El capítulo dos se centra en el marco referencial con conceptos de redes sociales, Social media y la interactividad como un factor para encontrar y calcular el User Engagement en Facebook.

El capítulo tres describe la implementación de las técnicas de Social media identificadas en el estudio de caso de la fan page de Facebook, donde se aplica una metodología de intervención y con los resultados de plantea un modelo integral que incorpora un enfoque de flujos de conocimiento identificando herramientas tecnológicas se presenta como apoyar a los procesos de Social media. A partir de este resultado, se explica la construcción de la propuesta de marco de trabajo para el minado de Social media basado en Gestión de Conocimiento (GC).

El capítulo cuatro por medio de experimentación con tres conjuntos de datos describe el análisis de las relaciones entre las métricas de Facebook con el cálculo de User Engagement, enfatizando la importancia de las métricas extraídas, adquiriendo diferentes modelos de regresión. Los datos preprocesados de Facebook permitieron obtener propuestas para un modelo de regresión con un mejor ajuste para calcular e interpretar el User Engagement.

Finalmente, el capítulo cinco presenta las conclusiones de este estudio de investigación, así como el trabajo futuro.

CAPÍTULO II: MARCO REFERENCIAL

El presente capítulo está dedicado a abordar los temas fundamentales como las investigaciones previas y los conceptos metodológicos. A través de este marco referencial se identificaron las oportunidades para la aportación de nuevo conocimiento.

2.1 Redes sociales

Las redes sociales son plataformas que tienen como objetivo permitir que los usuarios se relacionen, se comuniquen, compartan contenido, dirijan su aprendizaje y tengan presencia en internet. Estas plataformas se han convertido en canales esenciales para la transmisión de información y la comunicación, la formulación de proyectos, la concepción de comunidades, nuevos esquemas para entretenimiento e interacción social colectiva (Ureña et al., 2019). En este contexto, una red social se configura como una herramienta que transforma a las personas en consumidores y generadores de contenido (Gonzalez & Munoz, 2016). Según Celaya (2008), existen tres clasificaciones principales de redes sociales:

- Redes profesionales (LinkedIn, Xing, ResearchGate).
- Redes generalistas (Facebook, Twitter, Tuenti, Hi5).
- Redes especializadas (Instagram, eBuga, CinemaVIP, 11870).

Las redes sociales profesionales han contribuido a promover el concepto de networking entre las pequeñas y medianas empresas y mandos medios en el caso de las grandes empresas. Su valor agregado es que permiten desarrollar una amplia lista de contactos profesionales, tanto para intercambios comerciales e interacción, como para la búsqueda de oportunidades entre personas (Celaya, 2008). Normalmente, los usuarios que ingresan a estas redes incluyen una descripción de su currículum, que es el punto de partida en sus interacciones posteriores. Todas estas plataformas cuentan con un buscador interno que nos permite rastrear la base de contactos por el nombre de la persona u organización que estamos buscando. Los usuarios a su vez pueden integrarse según sus expectativas, gustos y preferencias (Herrera, 2012).

Herrera (2012), expuso que ha sido la aparición de las redes sociales especializadas en una determinada actividad social o económica, un deporte o un tema un punto importante de

incremento en las redes sociales. Esto aborda la inclinación natural de las personas a unirse a comunidades que comparten características e intereses similares.

Las redes sociales generalistas son espacios que cuentan con perfiles de usuarios que se integran según sus expectativas, gustos y preferencias, para comunicarse o compartir música, videos, fotografías e información personal (Celaya, 2008). En el contexto educativo, la aparición de las redes sociales hace que los procesos de enseñanza-aprendizaje tengan un carácter más social y personalizable en los estudiantes, incorporando las redes sociales en las actividades académicas (Gil-quintana et al., 2022; Prestridge et al., 2021; Shang et al., 2011). Los estudiantes acceden a contenido de destacados profesores e investigadores de todo el mundo a través de blogs, wikis, videos en línea, podcasts y recursos educativos abiertos (Orioque, 2021; Sánchez-Rodríguez et al., 2015) para apoyar actividades académicas y sociales (Gonzalez & Munoz, 2016). Un ejemplo de redes sociales de tipo generalista es Facebook, donde interactúan 2.9 billones de usuarios según el informe enero 2020 (Shaulova, 2020).

2.2 Social media

Los Social media son presentados como sistemas abiertos a todo el mundo, donde se fomenta la participación y se difumina la línea entre comunicador y audiencia, generando con esto una conversación multidireccional, con comunidades de gente interesada en los mismos temas (Martín Blanco, 2015). La participación en las redes sociales con criterios de Social media, tiene que contener una cultura colaborativa y con grados de participación, definiendo los cursos de acción y diferenciando las tareas de las distintas áreas. Así como también, considerar las nuevas tecnologías en la fijación de las reglas y políticas organizacionales (Mazza & Palermo, 2018; Alles, 2012).

El término Social media, se emplea para referirnos a aplicaciones basadas en Internet que utilizan los fundamentos ideológicos y tecnológicos de la Web 2.0, donde la Web 2.0 significa que el contenido y las aplicaciones ya no son creadas y publicadas por individuos, sino que se continúan modificados por todos los usuarios de manera participativa y colaborativa. Por ello, la búsqueda de las organizaciones por invertir en tecnologías Web 2.0, tienen como propósito

principal mejorar la interacción con sus clientes. En plataformas como Facebook o Twitter, las organizaciones se han convertido en un actor dentro de las relaciones (Goyal & Goyal, 2020).

2.2.1 Características

Existen una serie de factores transformando el modelo de comunicación entre las organizaciones, como la gran cantidad de comerciales, el incremento en las exigencias de calidad en la información, los influencers como líderes de opinión, las comunidades que empiezan a derivar de las nuevas tecnologías y la pérdida de credibilidad en los medios tradicionales (Humberstone, 2019; Shang et al., 2011). Las aplicaciones habituales de los Social media abarcan diversas actividades, tales como la difusión de eventos, visualizar discursos o supervisar las reacciones públicas ante anuncios, el análisis de comentarios e iniciativas, la identificación de opiniones e ideas respecto al comportamiento de una comunidad, descubrir eventos emergentes o viralización de contenidos y la realización de encuestas en grupos de redes sociales (Al-Maatouk et al., 2020; Batrinca & Treleaven, 2015). Las áreas estratégicas que deben monitorearse en las redes sociales son los términos de marca, las necesidades del usuario, los sentimientos o emociones y la comunicación. Este seguimiento es posible realizarlo directamente en redes sociales como Facebook (Kurnia & Suharjito, 2018). Las redes sociales están adquiriendo cada vez más importancia como canal para recopilar información sobre productos y servicios y aprovechar nuevas oportunidades (Penttinen & Frösén, 2022; Verhoef & Lemon, 2013).

Según el autor Ghani et al. (2019), en el caso del análisis de Social media es importante señalar que existen cuatro tipos de características: descriptiva, diagnóstica, predictiva y prescriptiva. Cuando se inicia un proceso de procesamiento de datos, la característica del análisis descriptivo es la fase inicial que proporciona todos los datos históricos necesarios. Además, se utiliza para facilitar el análisis de datos adicionales mediante una comprensión detallada de las partes exitosas y fallidas de los datos. Es implementada principalmente en organizaciones para informar eventos, como el seguimiento de ventas, departamentos y finanzas. Los modelos obtenidos a partir de una caracterización descriptiva permiten cuantificar, identificar y categorizar diversas relaciones en los datos (Simpao et al., 2014).

En el caso de la característica de diagnóstico, se le considera un tipo mejorado de análisis con la cualidad de trabajar con procesos como el descubrimiento de datos, la estadística, la

extracción de datos y las correlaciones de datos. En el diagnóstico se analizan los datos y se ordenan para lograr responder preguntas como acerca de por qué sucede algún evento. En este sentido, se investiga exhaustivamente los datos para conocer el comportamiento detallado y las causas de los eventos (Wang et al., 2016). El diagnóstico brinda la oportunidad de comprender los datos y responder rápidamente a algunas preguntas críticas sobre los datos involucrados. Por otro lado, en el predictivo se transforman los datos en información valiosa y procesable (Ayhan et al., 2013). Esta característica utiliza los datos para decidir el posible resultado futuro de un evento o la posibilidad de que ocurra una determinada situación. Por lo tanto, implica usar diferentes métodos estadísticos que van desde el modelado, el aprendizaje automático y la teoría de conjuntos para analizar hechos presentes y pasados, y así pronosticar eventos futuros. Los datos históricos que se encuentran en los modelos predictivos se utilizan en las organizaciones para identificar riesgos y oportunidades. El modelo resultante debe ser capaz de identificar las relaciones entre varios factores para permitir la evaluación de riesgos basada en un conjunto específico de condiciones, que guiarán el proceso de toma de decisiones. De manera similar, el análisis prescriptivo propone oportunidades de decisión para explotar una oportunidad futura o disminuir riesgos futuros. Este análisis demuestra los efectos de cada oportunidad de decisión. En la práctica, el análisis prescriptivo procesa constante y espontáneamente nuevos datos para mejorar la precisión y ofrecer buenas opciones de decisión (Ghani et al., 2019). Un enfoque prescriptivo permite examinar las posibles decisiones, las conexiones entre las decisiones, los impactos y resultados, que finalmente proponen una opción ideal en tiempo real.

2.2.2 Técnicas

En el análisis de Social media se requiere de técnicas para la recopilación, el seguimiento y el análisis de datos (Elkaseh et al., 2016). En el proceso de recopilación de los datos, se obtiene un patrón adecuado para el análisis de datos durante las conversaciones e interacciones. De esta forma, el análisis por Social media implica el uso de diferentes técnicas analíticas y de modelado de diversos campos. Además, las representaciones gráficas de redes sociales se utilizan en el análisis de Social media para proporcionar una comprensión detallada de su estructura fundamental, conexiones y propiedades teóricas (Al-Maatouk et al., 2020).

Grafos

Para extraer la información representada en los grafos, se necesita a) definir métricas que describan la estructura global de los grafos, b) encontrar la estructura comunitaria de la red, y c) definir métricas que caractericen los patrones de interacción local en los grafos (Biswal & Muley, 2018). Por lo tanto, una red social de características unidimensionales S se puede modelar como un grafo no dirigido $G = (V, E)$, donde $V = \{1, \dots, n\}$ es el conjunto de vértices o nodos que representan a los usuarios o entidades en S y $E = (E_{ij}), i, j \in V$ es el conjunto de aristas que conectan elementos de $V \cdot n = |V|$ representando el número de nodos en G y $m = |E|$ denota el número de aristas en G .

De tal forma que un grafo G puede representarse como una matriz de adyacencia $n \times n$ en $A = (A_{ij})$, para $i, j \in V$, donde:

$$A_{ij} = \begin{cases} w_{ij} \text{ para una arista entre } i \text{ y } j \text{ con peso } w_{ij}, \\ \mathbf{0} \text{ de lo contrario} \end{cases} \quad (1)$$

El concepto de grafo y la idea de una comunidad es una construcción integradora cuyo propósito es estudiar las redes sociales para comprender el comportamiento de los elementos que componen una comunidad, a través de las medidas o métricas. Las métricas permiten hacer una caracterización de los nodos en la red y las relaciones existen entre ellos (Kuz et al., 2016). Las métricas más comunes para entender los grafos y los nodos, permiten determinar la importancia y el rol de un nodo. Las más usadas se clasifican en métricas de poder y centralidad, y métricas de grupos (Hanneman & Riddle, 2009). En la herramienta Gephi las métricas utilizadas incluyen la distancia media de nodo a nodo, la Centralidad de Intermediación, la cual es un indicador de la centralidad de la red; la Centralidad de cercanía que describe los nodos que a pesar de tener pocas conexiones, sus arcos permiten llegar a todos los puntos de la red más rápidamente que desde cualquier otro punto y la detección de comunidades, como una medida de la estructura de las redes, diseñada para a través de una red en módulos medir la fuerza (también llamados grupos o comunidades) (Kuz et al., 2016).

La herramienta Gephi tiene 12 algoritmos que se pueden seleccionar para distribuir los nodos de un grafo (Humberstone, 2019). Los más importantes dentro de esta técnica son los algoritmos que simulan una diversa cantidad de fuerzas entre los nodos para modificar su posición,

conocidos como algoritmos guiados por fuerzas (Tabla 1), por lo que, los nodos conectados se atraen entre sí y los nodos no enlazados se separan o aíslan (Diliana & Indrawati, 2022).

Tabla 1 Algoritmos de distribución seleccionados en Gephi (Bedi & Sharma, 2016; Diliana & Indrawati, 2022; Humberstone, 2019).

Algoritmo	Descripción
Fruchterman-Reingold	Se trata de una distribución que organiza todos los nodos formando una circunferencia, resultando útil para obtener una representación visual de cada uno de los nodos presentes en la red.
Yifan hu	Es un algoritmo que se elige para reducir la complejidad del cálculo debido a su característica multinivel, y el grado de nodo se usa para representar el color y el tamaño de un nodo.
Noverlap	Es un tipo de algoritmo que tiene la función de apoyar a los que no tienen incorporado nativamente el parámetro de evitar la superposición entre nodos (por ejemplo: Yifan Hu Multilevel u OpenOrd).
Expansión	Esta distribución permite expandir la red sin cambiar el diseño que ya posee.
Ajuste de etiquetas	Cada nodo simboliza un conjunto de datos, como un tweet, un usuario o un hashtag. El algoritmo asegura que no exista superposición entre estas etiquetas, facilitando la comprensión del contenido de cada nodo.

Análisis de redes sociales (texto, opciones y sentimientos)

En las redes sociales se produce contenido en función a la cantidad de usuarios activos, en este sentido las organizaciones están motivadas a comprender qué problemas y tendencias evolucionan para identificar riesgos y oportunidades en la comunicación y derivar implicaciones de beneficio. Por ello, el análisis de redes sociales es definido como la técnica de Social media que busca ampliar, adaptar y combinar métodos para el análisis de los datos provenientes de las redes sociales (Zeng et al., 2010). Entre los métodos más implementados se encuentran: la minería de textos y el análisis de opiniones y sentimientos.

En el método de Minería de textos (Text Mining) se aplica un proceso automatizado para detectar y revelar conocimientos, interrelaciones, patrones nuevos y descubiertos en recursos de datos textuales no estructurados. La minería de textos se centra en conocimientos no descubiertos dentro de grandes cantidades de texto. Este campo de investigación utiliza algoritmos de minería de datos, como clasificación, agrupamiento y reglas de asociación, para explorar y descubrir nueva información y relaciones en fuentes textuales. Es un campo de investigación interdisciplinario que combina recuperación de información, minería de datos, aprendizaje automático, estadística y

lingüística computacional (Younis, 2015). Por otro lado, en el análisis de sentimientos, también conocido como minería de opiniones, se enfoca en el análisis de subjetividad. Es el proceso mediante el cual es posible determinar la actitud o polaridad de opiniones o reseñas escritas por humanos para calificar productos o servicios. El análisis de sentimientos se aplica a cualquier forma textual de opiniones, como blogs, reseñas y microblogs (Younis, 2015; Milstein et al., 2008), se puede realizar analizando a nivel de documento o de oración. En el primer caso, se evalúa todo el documento para determinar la polaridad de opinión, donde primero se deben extraer las características que describen el producto/servicio. Mientras que, en el segundo, el documento se divide en frases cada una de las cuales se evalúa por separado para determinar la polaridad de opinión.

2.2.3 Procesos

Implementar argumentos basados en las interacciones de reacciones de los usuarios refleja la fuerza de la conexión social en un contexto de comportamiento colectivo, proporcionando perspectivas sobre cómo los usuarios de las redes sociales se comportan y son influenciados socialmente (Meng et al., 2020). En el desarrollo de modelos analíticos, es crucial que los gestores de redes sociales se apoyen de estrategias que fomenten la interacción en las publicaciones, con el fin de incrementar el compromiso de los usuarios y por ende la visibilidad en internet (Peruta & Shields, 2017). En lo que respecta a comprender los modelos de análisis de Social media, se han documentado los procesos principales para Social media (Figura 2): el descubrir los datos (Mushtaq et al., 2018), la recopilación y preparado de datos y el análisis de datos (Goyal & Goyal, 2020; Gupta & Gupta, 2016).



Figura 2 Procesos de Social media. Adaptado de: Goyal & Goyal (2020); Meng et al. (2020); Mushtaq et al. (2018); Gupta & Gupta (2016); Liu et al. (2016).

Los estudios realizados por Goyal & Goyal (2020), Meng et al. (2020) y Liu et al. (2016) respaldan el uso de cuatro actividades fundamentales para el proceso de análisis de datos, al mismo tiempo que incorporan dentro del ambiente del Big Data los datos generados por las redes sociales. Además, destacan la importancia crucial de esta etapa de análisis, ya que la mayoría de los desafíos para los investigadores se presentan en este punto (Liu et al., 2016). Estas propuestas de modelos analíticos abarcan las actividades y los flujos de información, modelos de roles y/o acciones, así como modelos de flujo de trabajo.

2.3 Algoritmo en un enfoque de minería de datos

Los algoritmos siempre han sido una parte esencial de la vida profesional, personal y social; en todas las verticales e industrias, han facilitado el desarrollo de las ciencias naturales, sociales, humanidades y han permitido que la Tecnología logre enriquecer el conocimiento (Abiteboul & Doweck, 2020). Los algoritmos de las redes sociales son un ejemplo de la necesidad de aplicar conjuntos ordenados y finitos de operaciones que permitan hallar la solución a una problemática. Sin embargo, como se muestra en la Tabla 2 dependiendo del área de aplicación, el enfoque y el concepto que adaptan, puede variar.

Tabla 2 Definiciones de algoritmo para diversas áreas de aplicación.

Autor(es)	Concepto	Área de aplicación
Liao, S. H., Chu, P. H. y Hsiao, P Y., 2012	Es un proceso detallado para llevar a cabo un cálculo. La estructura de un algoritmo se describe como una lista limitada de instrucciones claramente definidas para realizar una función específica. Estos son empleados para realizar cálculos, procesar datos y realizar razonamiento automático.	Minería de datos
Hill, R. K., 2015	Se trata de una estructura de control compuesta finita, abstracta, efectiva, presentada de manera imperativa, que cumple un propósito específico bajo ciertas condiciones. En contraste con los enfoques tradicionales, las definiciones recursivas se descartan debido al énfasis en la imperatividad.	Informática
Seaver, N., 2019	Se hace referencia a cualquier procedimiento computacional bien definido y finito que toma uno o más valores como entrada y produce uno o más valores como salida. Por ende, un algoritmo es una serie de pasos computacionales que transforman la entrada en la salida.	Computación
Erciyas, K., 2021	Consiste en un conjunto limitado de instrucciones o lógica, dispuestas en orden, destinadas a llevar a cabo una tarea específica predefinida. Cualquier algoritmo debe cumplir con ciertos requisitos; puede o no aceptar entradas, pero se espera que proporcione alguna forma de salida, que constituye la solución al problema en cuestión.	Matemáticas y teoría de grafos

2.4 Interactividad

La interactividad de forma general en las redes sociales, y particularmente en Facebook, apoya en generar una mayor actividad de los usuarios hacia la marca, producto o servicio (Kaye, 2021). Según el reporte emitido en enero de 2022 por el portal Statista, clasificando el número de usuarios activos mensuales a diversas redes sociales, Facebook a diecinueve años de su creación continúa siendo considerada la red social más popular en México (Statista Portal, 2023). Desde un punto de vista tecnológico, la interactividad es la medida de cuánto participan los usuarios en la modificación del formato y contenido de un entorno en tiempo real. El objetivo de entender la interactividad como estrategia es aprovechar los datos del seguimiento y análisis de las publicaciones en las redes sociales, beneficiándose del conocimiento experto (Steinmetz et al., 2020). La interactividad se mide, analiza y discute a partir de métricas de interacción y conversación, como reacciones, compartidos y comentarios, que son las más comunes en Facebook (Furtado Avanza & Moreira Pinheiro, 2018; Sutanto, 2016). La interactividad da forma a la manera en que los usuarios perciben la información, así como a las futuras decisiones de los algoritmos

sobre sus sugerencias y preferencias. Se asume que un alto nivel de Interactividad generará diversos tipos de actitudes (positivas y favorables o negativas) sobre la fan page o grupos explorados en Facebook (Kaye, 2021). Las ventajas de Facebook están relacionadas con el acceso, manejo e intercambio de información en la red social (Al-Dheleai & Tasir, 2017).

2.5 User Engagement en Facebook

En Facebook el User Engagement es un valor cuantitativo de la interacción bidireccional entre los usuarios y las organizaciones en las redes sociales (Oviedo-García et al., 2014), se enfoca en conocer por medio del comportamiento de los usuarios la interactividad en las redes sociales (Froment et al., 2022). Para ello, se sugiere usar una serie de métricas, cuantificando variables relacionadas con la interacción con el contenido, el número de reacciones, el intercambio de publicaciones y comentarios (Molina et al., 2021; Savelev et al., 2021; Kim & Yang, 2017; Chen et al., 2013). Los trabajos recientes señalan que debe preferir en el análisis de User Engagement tanto un punto de vista positivo o negativo (Shahbaznezhad et al., 2021; Ballesteros Herencia, 2018), donde los factores que influyen en su cálculo incluyen: la cantidad de me gusta o reacciones (Molina et al., 2021; Savelev et al., 2021; Chen et al., 2013), los comentarios, las acciones, el tipo de contenido o los días en que se hace la publicación (Velazquez-Solis et al., 2022; Chan-Olmsted et al., 2017).

2.5.1 Características

La medición del User Engagement se calcula mediante tres métodos diferentes: encuestas y entrevistas, medidas implícitas y análisis web (Chan-Olmsted et al., 2017). En la analítica web, el User Engagement se calcula a través de la actividad en las plataformas digitales que generan los usuarios. Por lo tanto, se requieren una serie de métricas para cuantificar las variables relacionadas con la interacción, como el número de visitas, el tiempo dedicado, el número de seguidores o los clics del tipo de reacción (me gusta, me encanta, jaja, guau, enojado), el intercambio de publicaciones y los comentarios. (Jayasingh, 2019; Kim & Yang, 2017; Peruta & Shields, 2017; Hu & Chen, 2016). La Tabla 3 presenta diez fórmulas para calcular el User Engagement en función de los datos obtenidos de las páginas de fans o grupos de Facebook identificados durante la última década. Al revisar cada uno de los trabajos, ocho de ellos no tenían una justificación identificable de las variables ni verificación matemática sobre el origen de la fórmula que estaban

implementando. Tener un índice de las variables implementadas por diferentes fórmulas ayuda a identificar una medida integrada de compromiso, lo que subraya la importancia de los parámetros de una fórmula más normalizada (Eriksson et al., 2019). En este trabajo, las fórmulas propuestas por Eriksson et al. (2019), Jayasingh (2019) y Ge y Gretzel (2017) fueron seleccionados por demostrar mejores resultados en el cálculo de Engagement al implementar coeficientes. Solo el trabajo de Jayasingh (2019) identificó que no todos los seguidores interactúan con el contenido (representado por el exponente 0.8) y que el esfuerzo necesario para compartir o comentar una publicación es mayor, representado por los coeficientes 5 y 10, representando una ventaja sobre las otras fórmulas.

Tabla 3 Fórmulas para calcular el Engagement en Facebook.

Autores	Fórmula propuesta
Bonsón & Ratkai, 2013 Hoffmann, 2013	$\frac{N_L + N_C + N_S}{\frac{N_{Messages}}{N_F}}$
Niciporuc, 2014	$\frac{N_L + N_C + N_S}{N_F}$
Oviedo-García et al., 2014	$\frac{\frac{N_L + N_C + N_S}{N_F}}{\mu_{Scope}}$
Herrera-Torres et al., 2017	$\frac{N_R + N_S + N_C}{N_F}$
Ge & Grezel, 2017 Eriksson et al., 2019	$N_R + 5N_S + 10N_C$
Vadivu & Neelamalar, 2015 Peruta & Shields, 2017	$\frac{N_R + N_S + N_C}{N_{likes\ fanpage}}$
Phuntusil & Limpiyakorn, 2017	$\frac{N_R + 5N_S + 10N_C}{N_{Likes\ fanpage}^{0.8}}$
Ballesteros-Herencia, 2018	$\frac{N_R + N_S + N_C}{Scope\ post}$
Jayasingh Sudarsan, 2019	$\frac{N_R + 5N_S + 10N_C}{N_{Likes\ fanpage}^{0.8}}$
Martínez-Sala & Segarra-Saavedra (2020)	$\frac{N_R + N_S + N_C}{N_{Post} \times N_{Fans}}$

Abreviaciones: N_L = Número de likes, N_R = Número de reacciones, N_S = Número de veces que es compartido, N_C = Número de comentarios, N_F = Número de seguidores en página de fans o grupo en Facebook.

2.6 Identificación de las áreas de aplicación de Social media apoyadas en modelos, técnicas y herramientas

De acuerdo con lo antes expuesto, resulta fundamental examinar las estrategias de Social media como componente integral de un marco de referencia para el análisis de información extraída de las redes sociales. Se optó por emplear un mapeo sistemático de la literatura, según la metodología propuesta por Petersen et al. (2015), con el propósito de identificar las herramientas usadas, las técnicas y los modelos de análisis de Social media que han sido aplicadas en el análisis de redes sociales y que han sido propuestas en diversas áreas de aplicación. El procedimiento consta de cinco fases para obtener un panorama completo en el ámbito abordado (Brenes Carranza et al., 2020; Ramírez-Alpízar et al., 2020).

El primero consta en formular las preguntas de investigación, y con el propósito de que se cumpla el objetivo se construyó una cadena de búsqueda exploratoria para guiar el proceso. El segundo, determina una estrategia de búsqueda y el proceso de selección de estudios en función a los criterios de inclusión y exclusión. Para ello, se construyó una cadena de búsqueda en función a los términos clave y se seleccionaron las bases de datos. En este paso se recuperaron de cada base de datos: 28 en IEL, 46 en Scopus, 13 en Springer Link, 222 en Elsevier y 11,100 en Google Scholar, considerando AMC y Wiley (Figura 3).

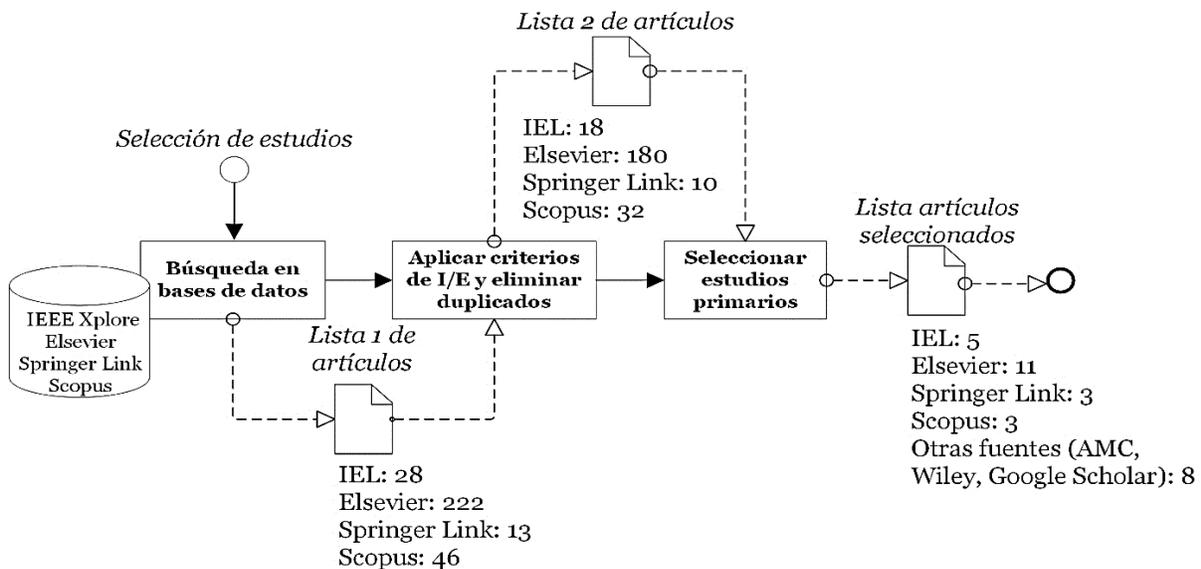


Figura 3 Proceso de Selección de artículos.

El tercer paso, es una evaluación de la calidad de los estudios para de acuerdo al indicador de calidad, seleccionar los estudios primarios; El cuarto consiste en presentar los desafíos al valor del estudio, así como las maneras de minimizar el impacto; Y finalmente, el proceso de extracción y análisis permitió identificar la información relevante y agrupar los estudios primarios en cuatro categorías: 1) Enfoques, 2) Modelos de análisis para Social media, 3) Arquitecturas de software basadas en Social media y 4) Técnicas/Herramientas de Social media. Así mismo, se asignó una categoría a cada artículo en función de su área de aplicación: Gestión de negocios, Mercadotecnia, Educación e Investigación. Estas categorías específicas permitieron identificar aquellos estudios que respaldaban sus propuestas mediante investigaciones empíricas y que analizaban la aplicación de modelos, herramientas y técnicas.

2.6.1 Identificación de modelos para Social media en diversas perspectivas

Se identificaron los artículos que se centran en un marco conceptual orientado a Social media o que presentan la implementación de escenarios de prueba o estudios de caso. En la Tabla 4, se resaltan 8 artículos, representando el 26.67% del total de estudios primarios. De estos, 5 están vinculados al ámbito de Investigación y 3 al ámbito de Educación, lo que demuestra el interés en ofrecer definiciones que faciliten la interpretación y adaptación de los conceptos asociados a Social media. Autores como Han et al. (2020), Stieglitz et al. (2018) y Gupta y Gupta (2016) proporcionan revisiones de conceptos, abordando temas como Análisis de sentimientos (Sentiment Analysis), Analítica de Social media (Social media analytics) y Minería de opiniones (Opinion Mining).

Tabla 4 Estudios con áreas de identificación con orientación a Social media.

Área de aplicación	Estudios con enfoques de Social media	Cantidad
Educación	Buxarrais, M. R. (2016), Amin & Rajadurai (2018), Chugh & Ruhi (2018)	3
Investigación	Gupta & Gupta (2016), Stieglitz et al. (2018), Goyal & Goyal (2020), Han et al. (2020), Meng et al. (2020)	5

Se enfatiza la relevancia de analizar los datos recopilados de las redes sociales, ya que la mayoría de los desafíos para los investigadores surgen durante esta fase (Liu et al., 2016). En la Tabla 5 se presentan 11 artículos seleccionados clasificados por áreas de aplicación en la categoría de modelos de análisis de Social media. Es significativo señalar que el 64% de estos estudios se

centra en modelos de análisis social, provenientes de investigaciones empíricas en lugar de modelos teóricos (Velazquez-Solis et al., 2021).

Tabla 5 Estudios con Modelos de análisis basados en Social media.

Área de aplicación	Estudios con modelos de análisis de Social media	Cantidad
Educación	Peruta & Shields (2017) - Estrategias y participación por escuela Salloum et al. (2017) - Modelo de Procesamiento de Lenguaje Natural He et al. (2017) - Modelos de temas, Análisis discriminante lineal (LDA) y Análisis semántico latente probabilístico (PLSA). Tocoglu et al. (2019) - Modelo de emociones básicas de Paul Ekman, Red Neuronal Artificial (RNA), Red Neuronal Convolutacional (RNC) y Red Neuronal Recurrente (RNR).	4
Mercadotecnia	Hu & Chen (2016) - Modelo de utilidad basado en modelos de predicción Arora et al. (2019) - Regresión KNN y Regresión de Vectores de Soporte Bitiktas & Tuna (2020) - Motivación psicológica del boca a boca	3
Gestión de negocios	Kurnia & Suharjito (2018) – Naive Bayes, SVM y árboles de decisión Purohit, Castillo & Pandey (2020) – modelo de servicio	2
Investigación	Lipizzi et al. (2016) – Clústeres semánticos y Redes de relaciones Hwa Choi, S. (2019) – Modelado de sentimiento y cuantificación de índices de sentimiento utilizados en Social Big Board	2

2.6.2 Técnicas aplicadas y herramientas implementadas en Social media

Separado en dos partes, donde la primera presenta propuestas con Arquitecturas de software basadas en Social media que responden a las Técnicas aplicadas y herramientas implementadas en Social media. Se identificaron 6 estudios primarios (20%) en su mayoría enfocados en el área de Investigación (Tabla 6). Las propuestas no solo incluyen los diseños arquitectónicos, también los requisitos tecnológicos de una arquitectura que se encuentra orientada a servicios que funcionan con análisis de emociones y opiniones.

Tabla 6 Estudios enfocados en Arquitecturas de software basadas en Social media.

Área de aplicación	Estudios de Arquitecturas de software basadas en Social media	Cantidad
Educación	Adinolfi et al. (2016) – Arquitectura orientada a componentes Raspopović et al. (2016) – Arquitectura orientada a servicios	2
Gestión de negocios	Immonen et al. (2015) – Arquitectura orientada a componentes	1
Investigación	Mushtaq et al. (2018) – Arquitectura en capas Hayat et al. (2019) – Arquitecturas de aprendizaje profundo Pelález, Martínez & Vargas (2019) – Arquitectura en capas	3

La clasificación de Técnicas aplicadas y herramientas implementadas en Social media se centra en descubrir los métodos o recursos utilizados para llevar a cabo los procesos de Social media. De igual manera, se busca categorizar las distintas herramientas disponibles que respaldan la ejecución de Social media en diversas redes sociales. En la Tabla 7 se detallan los Técnicas aplicadas y herramientas implementadas en Social media correspondientes a cada ámbito de aplicación. La mayoría de los métodos, adaptados desde la Inteligencia Artificial con Big data o Machine learning y Minería de datos, apuntan a que Social media apoya en las actividades de comprensión del comportamiento virtual y una adaptación de procesos organizacionales, especialmente durante las primeras etapas de proyectos (Han et al., 2020).

Tabla 7 Estudios de técnicas y herramientas basadas en Social media.

Área de aplicación	Técnicas aplicadas y herramientas implementadas en Social media	Cantidad
Educación	Shen & Kuo (2015) – Técnicas de análisis de tendencias con minería de datos y massive open online courses (MOOCs) Al-Maatouk et al. (2020) – Technology acceptance model (TAM) y Task/Technology Fit (TTF)	2
Mercadotecnia	Younis, E. (2015) – Package 'twitter'	1
Gestión de negocios	Arafeh et al. (2020) – Neo4j Graph Platform	1
Investigación	Noor & Haneef (2020) – Hama and Spark tools y Jaql	1

2.6.3 Análisis del conjunto de datos

Las implicaciones de investigación y prácticas se enfocan en contribuir al establecimiento de un marco de trabajo en el desarrollo de metodologías y técnicas apropiadas para aprovechar la información en Social media. La Figura 4 presenta un gráfico de dispersión con las categorías y áreas de aplicación ordenadas de manera cronológica. Se observa una concentración del 37% en la categoría 2 (Modelos de análisis de Social media), de los cuales el 36.37% se dirige hacia el área 1 (Educación). Se destaca un 40% donde en las categorías asignadas al estudio la mayor incidencia se encuentra en el área 4 (Investigación). Algunos estudios comparten características tanto en área de aplicación como en categoría, indicando coherencia en el interés por un tema específico entre diversos autores y centros de investigación. El 57% de los estudios pertenece a los últimos tres años, y se resalta que el 43.33% de los artículos son de los años 2019 y 2020, confirmando el constante interés en los enfoques de Social media como una temática vigente de investigación.

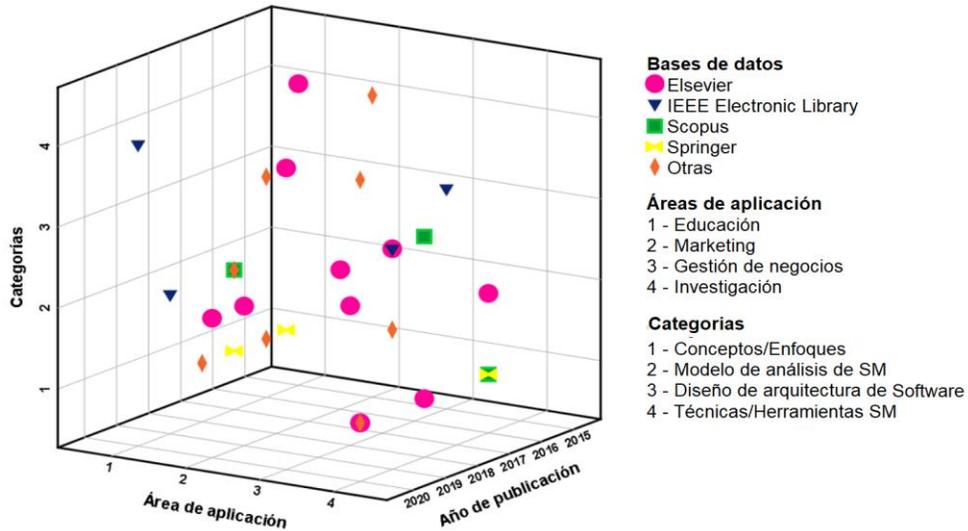


Figura 4 Gráfico de dispersión de puntos con la descripción de los estudios primarios.

Se analizó un grupo de 30 estudios primarios publicados entre 2015 y 2020, evaluando el panorama teórico y los resultados informados en prácticas y áreas de aplicación comunes, así como en las categorías establecidas. Se observó el beneficio derivado de la adopción de estas técnicas y herramientas, identificando un crecimiento en el interés de las organizaciones por incorporarse en sus procesos habituales. A pesar de estos avances, persisten desafíos en su adopción. Los estudios analizados aún se centran en un criterio de Minería de datos, orientado por el proceso de Knowledge Discovery in Databases. Esto señala una oportunidad para proponer marcos de trabajo y referencias metodológicas específicas para Social media. En este contexto, se evidencia la carencia de esquemas de trabajo que guíen y rastreen los artefactos y las actividades vinculadas a un proceso de Social media, ya que no consideran la perspectiva de los elementos de los procesos de software en su implementación.

CAPÍTULO III: MARCO DE TRABAJO PARA EL MINADO DE SOCIAL MEDIA BASADO EN GC

El capítulo presenta como se implementó una metodología de intervención para obtener un modelo integral que incorpora un enfoque de flujos de conocimiento identificando herramientas tecnológicas que apoyen en el análisis de los procesos de Social media. A partir de este resultado parcial, se explica la constitución de la propuesta de marco de trabajo para el minado de Social media basado en Gestión de Conocimiento (GC).

3.1 Metodología de intervención

Se aplicó una versión adaptada de la metodología de intervención propuesta por González et al. (2009) y la metodología de flujos de conocimiento desarrollada por Flores-Rios (2016), la cual se organiza en fases y actividades, incluyendo: 1) una Auditoría del conocimiento que comprende la identificación y modelado de fuentes y tipos de conocimiento, así como cuatro procesos de creación de conocimiento; 2) la identificación de herramientas tecnológicas asociadas a los procesos de creación de conocimiento que respaldan las actividades de Social media; 3) la implementación de un sistema de GC; y 4) el monitoreo y evaluación (Velazquez-Solis & Flores-rios, 2021). A través de la aplicación de esta metodología, se obtuvo información fundamental para los procesos, técnicas y herramientas relacionados con Social media (Velazquez-Solis & Flores-rios, 2021).

3.1.1 Auditoría de conocimiento

La fase de Auditoría de conocimiento facilitó la identificación de oportunidades, problemas o necesidades y estrategias de conocimiento, sirviendo como fundamento para la identificación de procedimientos, herramientas tecnológicas, sistemas de información, roles y documentos asociados a los procesos de Social media. La Figura 5 proporciona una representación visual mediante una técnica de gráfica rica adaptada, mostrando ejemplificaciones de los tipos de conocimiento que aplican los roles a nivel operativo y estratégico. Entre los roles propuestos por Budiño (2011) y Alfaro & Watson-Manheim (2015) se encuentran en un nivel operativo (Figura

5), el Analista de Social media y el Administrador o responsable de la comunidad en redes sociales (Social media Community Manager), y a nivel estratégico, el coordinador, el director y gerente de redes sociales. Se anticipa que estas funciones contribuyan a mejorar la participación de los usuarios, interactúen de manera proactiva con los seguidores, establezcan una presencia efectiva en las diferentes redes sociales y promuevan comunidades externas. Asimismo, se espera que impulsen iniciativas de alcance y generen mayor compromiso para involucrar a la comunidad de forma activa.

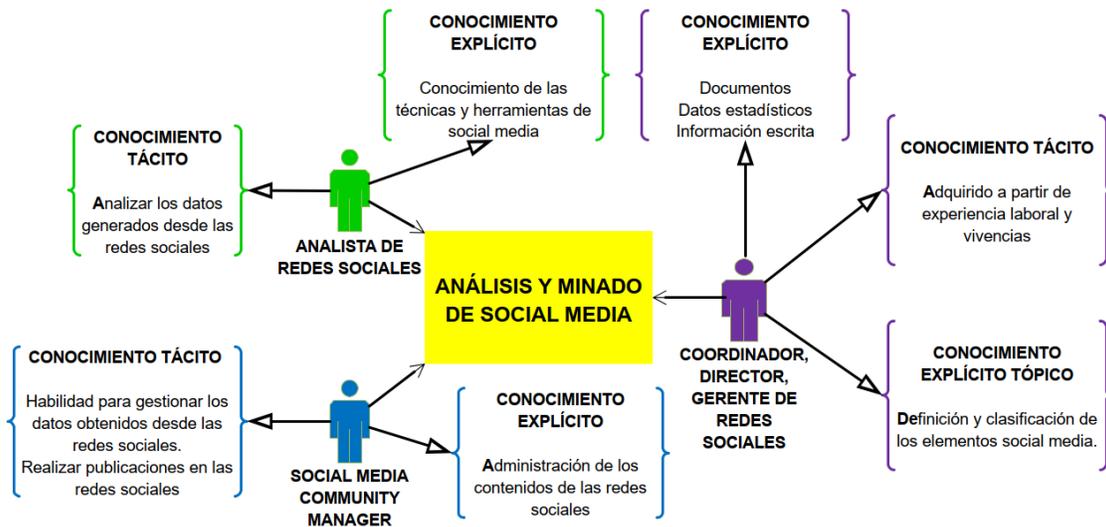


Figura 5 Identificación de roles de redes sociales y tipos de conocimiento.

Los roles identificados en el nivel operativo poseen conocimiento tácito que abarca las experiencias, el conocimiento y las habilidades necesarias para la implementación de técnicas en Social media. El analista de redes sociales otorga a la información extraída el sentido necesario para los datos generados por las redes sociales con su conocimiento tácito, al mismo tiempo que se genera y usa el conocimiento explícito relacionado con las técnicas y herramientas implementadas en Social media. Por otro lado, el Social media community manager debe poseer habilidades para realizar publicaciones, gestionar los datos generados por la interactividad de los usuarios, incluyendo la obtención del conjunto de datos para el analista de redes sociales. El conocimiento explícito generado por el Social media community manager y el analista de redes sociales se usa como punto de partida para establecer y diseñar las características de Social media aplicables en diversas estrategias.

3.1.2 Identificación de herramientas tecnológicas

Una parte fundamental fue detallar los elementos de los procesos vinculados a la generación de conocimiento, empleando un enfoque conocido como la espiral del conocimiento o modelo SECI (Nonaka & Takeuchi, 1995), que identificó las fuentes y categorías de conocimiento capturadas. Durante la fase de Auditoría de conocimiento, se evidenció que los responsables del área de TI (asumiendo roles centrados en Social media y Gerente de redes sociales) desempeñan funciones estratégicas y operativas, encargados de seleccionar y proporcionar las herramientas tecnológicas que respaldan los procesos de Social media, asegurando la continuidad de la información en concordancia con los cuatro procesos de creación de conocimiento. Estos hallazgos se reflejan en la conclusión de la fase 2, que especifica las tecnologías, incluyendo sistemas de información, herramientas tecnológicas y recursos de hardware, asociados con los cuatro procesos de conocimiento y que ofrecen apoyo a los procesos de Social media (Tabla 8).

Tabla 8 Asociación de tecnologías con Procesos de conocimiento desde la perspectiva de Social media.

Tipo de conocimiento	<i>A Tácito</i>	<i>A Explícito</i>
<i>De Tácito</i>	<p>Proceso: Socialización</p> <p><i>Característica:</i> Adquirir conocimiento mediante la colaboración, la identificación de expertos y la celebración de reuniones de equipo, ya sea de manera sincrónica o asincrónica.</p> <p><i>Tecnologías:</i> Facebook Llive, Twitter, Whatsapp, correos electrónicos, chat, video llamadas.</p> <p><i>Proceso de Social media:</i> descubrir los datos</p>	<p>Proceso: Exteriorización</p> <p><i>Característica:</i> Formular e intercambiar metáforas y analogías, narración de historias.</p> <p><i>Tecnologías:</i> Comentarios registrados en Facebook, vídeos, fotos. Foros, Talleres o seminarios, blogs, carteles o documentación de registro de experiencias.</p> <p><i>Procesos de Social media:</i> analizar los datos</p>
<i>De Explícito</i>	<p>Proceso: Interiorización</p> <p><i>Característica:</i> Adquirir conocimiento implícito mediante la formación, el entrenamiento, la lectura, el análisis y la revisión de documentos o artefactos.</p> <p><i>Tecnologías:</i> Herramientas que apoyan en la extracción de datos para su interpretación. Por ejemplo: Anaconda Python, MS-Excel, RStudio, SPSS.</p> <p><i>Procesos de Social media:</i> recopilar los datos y preparar los datos.</p>	<p>Proceso: Combinación</p> <p><i>Característica:</i> Poner a disposición el conocimiento claro y directo de la organización, utilizando taxonomías, resúmenes y categorización.</p> <p><i>Tecnologías:</i> Administradores de Bases de Datos. Por ejemplo: Gephi, MS-Excel, Anaconda Python, SPSS, Mongo DB, Neo4j, Matlab. Grafos, Gráficos, reportes estadísticos, tablas con resultados de información histórica, bases de datos, hojas de cálculo, Análisis estadístico de datos.</p> <p><i>Procesos de Social media:</i> preparar los datos y analizar los datos.</p>

3.1.3 Despliegue del sistema de GC

En la etapa de implementación, se espera que el sistema de Gestión del Conocimiento (GC) se aplique de acuerdo con las recomendaciones y estrategias definidas en los flujos de conocimiento. Por este motivo, se llevó a cabo un mapeo sistemático de la literatura para identificar áreas de aplicación que integraran los modelos de análisis, los enfoques, las herramientas y las técnicas de Social media (Velazquez-Solis et al., 2021). Este mapeo también destacó problemáticas asociadas a los flujos de conocimiento. Con el propósito de alcanzar un nivel de madurez significativo en la GC, siguiendo las pautas de Flores-Rios (2016) y retomando las indicaciones de Arias & Aristizábal (2011), es necesario llevar a cabo una etapa exploratoria abarcando áreas clave en la estrategia, los procesos de GC, las tecnologías de la información y la innovación. La Tabla 9 presenta una estructura de áreas estratégicas de flujos de conocimiento basada en el trabajo de Quiñonez Zuñiga & Rivera Martinez (2021) en el modelo de ayuda para la Gestión Integrada del Conocimiento (MAGIC).

Tabla 9 Ámbitos de actuación de los Social media en las áreas estratégicas.

Áreas estratégicas	Ámbito de actuación	Rol involucrado
Revisión de estrategias de Gestión de Conocimiento	Identificar las colaboraciones y los contenidos creados por los usuarios, para incentivar al debate de las temáticas centrales de la red social. Preparar al analista de redes sociales en métodos de recopilación de datos (Socialización, Interiorización).	Social media community manager Coordinador, Director y Gerente de redes sociales
Análisis de los aspectos clave	Realizar estudios con diferentes muestras extraídas de las interacciones de los usuarios (Combinación, Exteriorización).	Analista de redes sociales <i>Social media community manager</i>
Planes de innovación	Incluir en la planificación de procesos la automatización del monitoreo y la mejora continua de las interacciones en plataformas de redes sociales (Combinación, Internalización, Externalización).	Coordinador, Director y Gerente de redes sociales
Análisis de los modelos de los procesos	Detectar tanto el conocimiento tácito como el explícito necesario para llevar a cabo el análisis de los datos extraídos. Identificar a individuos con habilidades innovadoras tanto en el desarrollo de software como en la investigación de tendencias tecnológicas.	Analista de redes sociales <i>Social media community manager</i>
Identificación de problemas	Establecer canales de comunicación entre todas las partes involucradas, identificar deficiencias técnicas y de conocimiento (Socialización, Externalización, Internalización).	Analista de redes sociales <i>Social media community manager</i>

Benchmarking	Evaluar la ejecución del análisis de datos y el diseño arquitectónico (Combinación, Internalización).	Analista de redes sociales Coordinador, Director y Gerente de redes sociales
Sistema de vigilancia tecnológica	Vigilar continuamente las interacciones, visualizaciones y expansión de la red social (Ballesteros Herencia, 2018) (Combinación). Contar con las herramientas adecuadas para llevar a cabo la vigilancia tecnológica (Quiñonez Zuñiga & Rivera Martinez, 2021).	<i>Social media community manager</i>

En las propuestas de modelos de análisis y su desarrollo resulta crucial que los roles encargados de la gestión de redes sociales adopten estrategias que fomenten la participación en las publicaciones con el fin de optimizar la visibilidad y la interacción de los usuarios (Peruta & Shields, 2017).

3.1.4 Seguimiento y evaluación

Los procesos de seguimiento y evaluación son fundamentales para respaldar la mejora continua en la Gestión del Conocimiento (GC). Los resultados de esta fase son cruciales para tomar decisiones acerca de la implementación de nuevas iniciativas dirigidas a mejorar los flujos de conocimiento o introducir procesos innovadores. La Figura 6 presenta una visión integral de las relaciones y los elementos en un contexto de Social media, concentrándose en las etapas iniciales de un análisis de datos de redes sociales. Iniciando con el acceder a los datos creados por las actividades y comunicación de los usuarios en las redes sociales (expuestos en estadísticas semanales, mensuales o anuales), las organizaciones pueden obtener una comprensión más profunda de la información pertinente. Se aplican técnicas de Data mining, como árboles de decisión, Naive Bayes o clústeres semánticos, junto con herramientas diseñadas para la exploración y explotación de datos extraídos de las redes sociales, tales como las librerías de Python o R y sistemas de administración de bases de datos. También se tienen en cuenta los roles y procesos asociados con el Social media. Se sostiene que la aplicación de estas técnicas y herramientas desde una perspectiva de flujos de conocimiento contribuirá a analizar el comportamiento social virtual, proporcionando un primer enfoque a las necesidades de los usuarios y facilitando el diseño de modelos de predicción centrados en métricas de Social media.

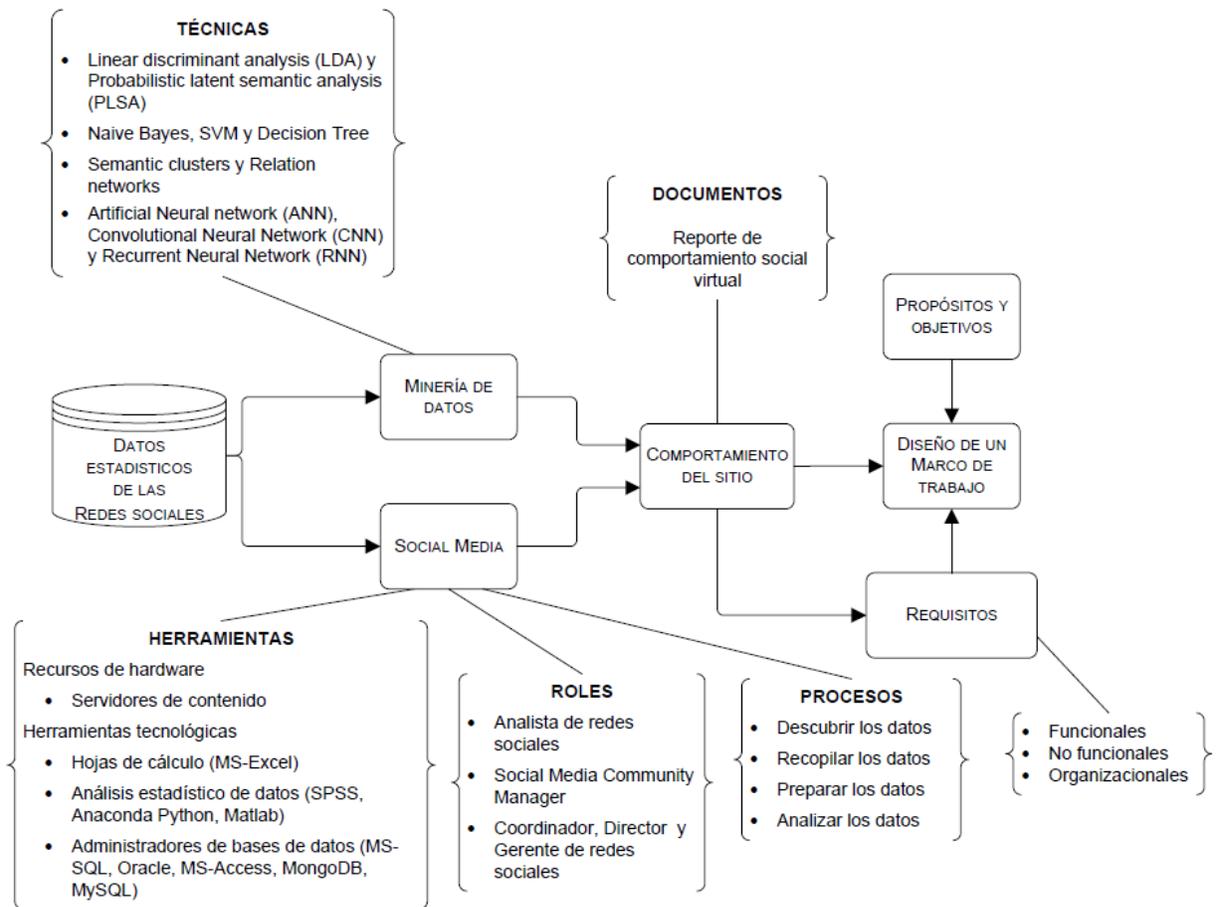


Figura 6 Modelo global para el diseño de un marco de trabajo basado en Social media.

3.2 Descripción del Marco de trabajo para el minado de Social media basado en GC

En la revisión de literatura se identificaron los desafíos y dificultades que involucran específicamente el proceso de Procesamiento de los datos cuando se encuentran bajo un enfoque de Social media. Esto representa que existe una carencia en expresar los procesos considerando las características específicas de los roles, las herramientas, las técnicas, los métodos y los resultados parciales de cada etapa que tienen los procesos como productos de trabajo. Por tal motivo, se propone el Social media Mining Framework como un marco de trabajo para la aplicación de procesos basados en técnicas y métodos de Social media con la definición de elementos básicos para un proceso software.

La Tabla 10 muestra los cuatro procesos definidos para el análisis por Social media: 1) Preparación y preprocesamiento de los datos, 2) Interpretación de los datos, 3) Análisis de datos (Ali et al., 2021; Pilař et al., 2021; Eriksson et al., 2019; El Haddaoui et al., 2018; Stieglitz et al., 2018; Bonzanini, 2016), y 4) Validación.

Tabla 10 Especificación de los tres procesos del Social media Mining Framework.

Nombre del Proceso	Propósito
1. Preparación y pre-procesamiento de los datos	Establecer las características para el análisis en el ambiente y dominio de conocimiento seleccionado (Ali et al., 2021; Eriksson et al., 2019; Stieglitz et al., 2018) para adquirir o recibir los datos a través de la selección de las métricas de interés (Bonzanini, 2016) y de esta forma realizar el preprocesamiento de los datos (limpiar, reducir e imputar, dar formato y transformar datos) (Gole & Tidke, 2015).
2. Implementación de técnicas y métodos	Aplicar las técnicas y los métodos de procesamiento de datos propios de los procesos de Social media (Eriksson et al., 2019; El Haddaoui et al., 2018) para interpretar y predecir comportamientos virtuales.
3. Análisis de los datos	Analizar el conocimiento descubierto y realizar una apropiada toma de decisiones (Stieglitz et al., 2018).
4. Validación	Verificar la diversidad del conocimiento descubierto comparando los resultados entre diferentes muestras de los datos.

Los elementos de cada proceso han sido caracterizados por un contexto de Gestión de Conocimiento y Análisis de datos. El Social media Mining Framework sugiere un enfoque de procesos y Gestión del Conocimiento, permitiendo la aplicación de métodos, técnicas y herramientas para el continuo análisis de los datos extraídos desde las redes sociales. La propuesta constituye un conjunto de buenas prácticas establecidas por estándares, guías y metodologías, que integra actividades fundamentales para el análisis, descubrimiento y verificación de conocimiento en redes sociales. Además, incluye técnicas para llevar a cabo las actividades, con el uso de herramientas propuestas y roles identificados. Los componentes que integran el Modelo de Referencia de Procesos presentado en la Figura 7 son:

- Propósito y objetivos
- Actividades
- Herramientas tecnológicas sugeridas
- Roles involucrados
- Productos de trabajo

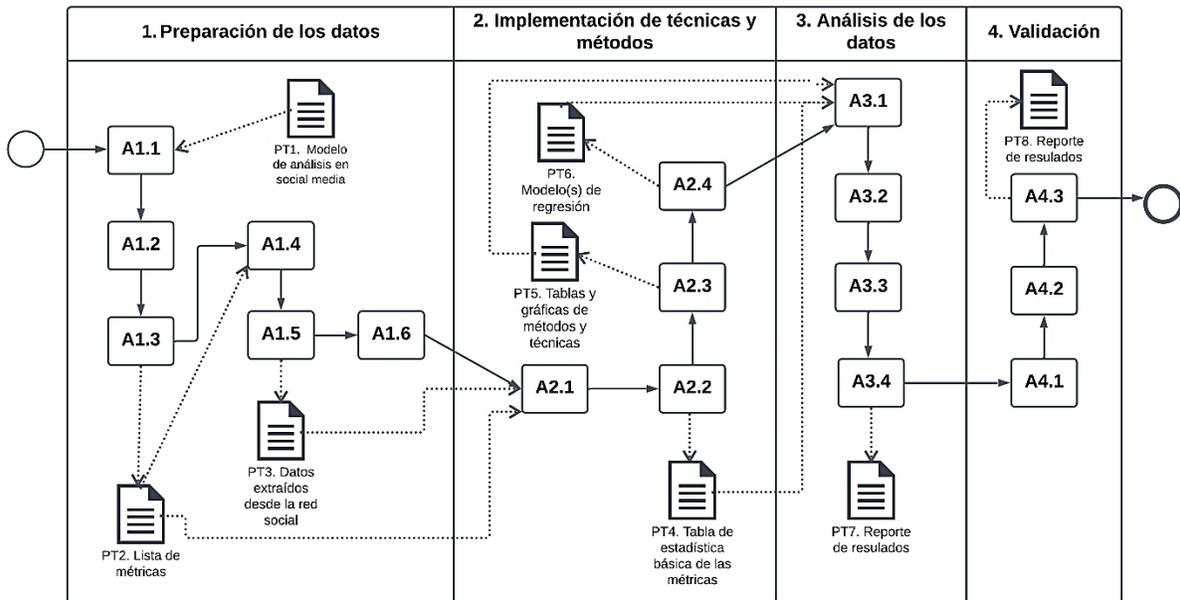


Figura 7 Flujo de trabajo de las actividades de los cuatro procesos de Social media Mining Framework.

3.2.1 Proceso 1: Preparación y preprocesamiento de los datos

En el capítulo II, la Figura 2, presenta los procesos de Social media identificados por los autores Goyal & Goyal, (2020); Meng et al., (2020); Mushtaq et al., (2018); Gupta & Gupta, (2016); Liu et al., (2016). En el proceso de Preparación de los datos se debe definir qué información en la red social se usará para responder a las preguntas planteadas con base en el ambiente establecido y el dominio de conocimiento de los roles participantes. Dado que esta actividad es primordial para obtener un conjunto de calidad, en el proceso se definió una estrategia para la sección de los estudios de caso, sujetos de estudio y periodos. El volumen del conjunto de datos fue un factor determinante para la selección de herramientas tecnológicas (Velazquez-Solis et al., 2022). Por lo tanto, para lograr establecer esta base inicial, se descargan los datos desde la red social y se hace el preprocesado de los datos.

1. Preparación de los datos	
<i>Propósito general</i>	
Establecer las características del análisis de los datos extraídos de la red social en el dominio de conocimiento y realizar el preprocesado de los datos.	
<i>Objetivos</i>	
O1.1 Identificar la fuente de datos de un periodo y el ambiente para que sea posible determinar las acciones en procesos posteriores. O1.2 Plantear qué información es de interés conocer y extraer desde la fuente de datos las métricas. O1.3 Seleccionar el algoritmo adecuado para calcular Engagement, interactividad o visibilidad.	
<i>Roles</i>	<i>Herramientas sugeridas</i>
<ul style="list-style-type: none"> — Social media community manager — Analista de redes sociales 	<ul style="list-style-type: none"> — Herramientas de redes sociales — Hojas de cálculo — Entornos de desarrollo integrado para análisis estadístico de datos
<i>Actividades</i>	
A1.1 Determinar un lenguaje en común para todos en el dominio del conocimiento y definir las preguntas que se van a responder con base en la red social seleccionada, el periodo de tiempo y el enfoque del ambiente.	
A1.2 Seleccionar el conjunto de variables requeridas por las métricas que se explorarán y analizarán.	
A1.3 Seleccionar el/los elementos de Social media que se analizarán en función a las métricas y algoritmos disponibles.	
A1.4 Obtener con alguna herramienta de redes sociales o con entornos de desarrollo integrado para análisis estadístico los datos desde la red social.	
A1.5 Realizar el preprocesado de datos, esto incluye calcular en una nueva columna el Engagement, interactividad o visibilidad, limpiar, imputar y transformar los datos para el análisis.	
<i>Productos de trabajo de entrada</i>	
<ul style="list-style-type: none"> — PT1. Modelo de análisis en Social media (Abdul Ghani et al., 2019) 	
<i>Productos de trabajo de salida</i>	
<ul style="list-style-type: none"> — PT2. Lista de métricas por página o por publicaciones. — PT3. Datos extraídos desde la red social. — PT4. Archivo con los datos preprocesados 	

En el proceso 1 se realiza no solo el preprocesamiento de los datos, también se hace una selección de las herramientas y técnicas que pueden implementarse para analizar la información. En algunos casos es necesario solucionar casos atípicos como los presentados en la Tabla 11.

Tabla 11 Casos atípicos en el preprocesado de los datos.

Casos atípicos	Solución propuesta
Reducir e imputar datos	<p><i>Duplicados de información:</i> Se recomienda codificar una rutina en el entorno de desarrollo integrado para análisis estadístico de datos que permita borrar los datos repetidos por tuplas.</p> <p><i>Datos faltantes:</i> Existen tres recomendaciones aplicables.</p> <ul style="list-style-type: none"> — La primera consiste en sustituir los datos faltantes por N/A o null, en el entendido de que los estadísticos básicos resultantes tendrán poca confiabilidad. — La segunda es sustituir los datos faltantes por ceros, pero en consecuencia los estadísticos básicos tendrían un sesgo hacia la izquierda y valores más bajos. — La tercera consiste en sustituir los datos faltantes por el cálculo del promedio de los datos que se tienen.
Dar formato a los datos	<p><i>Formato de la fecha:</i> Las métricas de fechas vienen en formato DD/MMM/AAAA HH:MM:SS AM/PM se recomienda elaborar una rutina que convierta las métricas de este formato en dos variables separadas.</p> <p><i>Categorización por etiquetas:</i> Existen métricas con formato categórico como el tipo de publicación o reacciones, se recomienda usar plataformas de software estadístico para cambiar el formato a uno numérico.</p>
Transformar los datos	<p><i>Formato incompatible en el software utilizado</i></p> <p>Se recomienda mantener formato xlsx en cualquier herramienta, puesto que el CVS en ocasiones causa problemas de importación.</p>

3.2.2 Proceso 2: Implementación de técnicas y métodos

En el proceso 2, se ejecutan las técnicas y los métodos de Social media, esto implica el aplicar métodos estadísticos como el modelado, teoría de conjuntos y minería de datos para interpretar y predecir comportamientos virtuales. Por consiguiente, los datos obtenidos desde la red social, en el formato indicado para la herramienta seleccionada son interpretados por diferentes características de Social media (descriptiva, diagnóstica, predictiva y prescriptiva) definidas en capítulo II.

2. Implementación de técnicas y métodos	
<i>Propósito general</i>	
Establecer una caracterización descriptiva que permita el cuantificar, identificar y categorizar las diversas relaciones en los datos según el dominio del conocimiento en un periodo determinado.	
<i>Objetivos</i>	
O2.1 Obtener una caracterización descriptiva de los datos O2.2 Identificar y categorizar las relaciones existentes entre los datos	
<i>Roles</i>	<i>Herramientas sugeridas</i>
— Analista de redes sociales	<ul style="list-style-type: none"> — Entornos de desarrollo integrado para el análisis estadístico de datos (RStudio, Anaconda Python) — Plataformas de software estadístico (IBM SPSS) — Entornos de Desarrollo Integrado para cómputo numérico (Matlab)
<i>Actividades</i>	
A2.1 Seleccionar las técnicas y métodos apropiados para responder las preguntas de A1.2 en función al periodo de tiempo (Tabla 12) A2.2 Obtener la relación de estadísticas básicas de las métricas seleccionadas en A1.2 A2.3 Aplicar métodos cualitativos de análisis estadístico y técnicas de Minería de datos A2.4 Realizar modelos de regresión para interpretar y/o predecir el comportamiento social virtual	
<i>Productos de trabajo de entrada</i>	
<ul style="list-style-type: none"> — PT1. Lista de métricas por página o por publicaciones. — PT2. Datos extraídos desde la red social. 	
<i>Productos de trabajo de salida</i>	
<ul style="list-style-type: none"> — PT4. Tabla de estadística básica de las métricas — PT5. Tablas y gráficas de métodos y técnicas — PT6. Modelo(s) de regresión 	

Tabla 12 Métodos, técnicas y herramientas sugeridas por cantidad de tuplas.

Cantidad de tuplas (Máximo)	Métodos cualitativos	Técnicas de minería de datos	Herramientas sugeridas
500 - 4999	<ul style="list-style-type: none"> - Estadística básica - Cálculo de correlaciones - Prueba de chi cuadrado 	<ul style="list-style-type: none"> - Detección de valores atípicos - Seguimiento de patrones - Gráfos 	<ul style="list-style-type: none"> - MS Excel - SPSS - Gephi
5000 - 14999	<ul style="list-style-type: none"> - Métodos para 500 – 4999 tuplas - Modelo de regresión para interpretación - Gráficas de Q-Q e histogramas 	<ul style="list-style-type: none"> - Patrones Secuenciales - Clustering - Gráfos 	<ul style="list-style-type: none"> - Anaconda - Python - RStudio - Gephi
15000 o más	<ul style="list-style-type: none"> - Métodos para 500 – 4999 tuplas - Métodos para 5000 – 14999 tuplas - Modelo de regresión para predicción 	<ul style="list-style-type: none"> - Clasificación - Clustering - Gráfos 	<ul style="list-style-type: none"> - Anaconda - Python - Gephi

3.2.3 Proceso 3: Análisis de los datos de redes sociales

El proceso 3, se convierte en el puente entre el análisis del conocimiento descubierto y la aplicación práctica de este entendimiento en la toma de decisiones. Al analizar tanto hechos presentes como pasados, se logra una visión integral que facilita la capacidad de prever eventos futuros, destacando la importancia de este proceso en la formulación de estrategias efectivas dentro del entorno de las redes sociales.

3. Análisis de los datos	
<i>Propósito general</i>	
Analizar el conocimiento descubierto para conocer el comportamiento detallado y las causas de los eventos y realizar una apropiada toma de decisiones.	
<i>Objetivos</i>	
O3.1 Identificar las relaciones entre las métricas para permitir la evaluación de factores. O3.2 Presentar los resultados del análisis a los roles destinados a la toma de decisiones.	
<i>Roles</i>	<i>Herramientas sugeridas</i>
<ul style="list-style-type: none"> — Analista de redes sociales — Coordinador de redes sociales — Gerente de redes sociales — Director de redes sociales 	<ul style="list-style-type: none"> — Entornos de desarrollo integrado para el análisis estadístico de datos (RStudio, Anaconda Python) — Plataformas de software estadístico (IBM SPSS) — Entornos de Desarrollo Integrado para cómputo numérico (Matlab)
<i>Actividades</i>	
A3.1 Realizar un análisis de los datos utilizando técnicas estadísticas y técnicas gráficas que pudieran ser univariadas, bivariadas o multivariadas.	
A3.2 Analizar las gráficas y tablas del PT4, PT5 y PT6 obtenidas del procesamiento de los datos de la red social.	
A3.3 Redactar un reporte de análisis detallado de los resultados	
<i>Productos de trabajo de entrada</i>	
<ul style="list-style-type: none"> — PT3. Tabla de estadística básica de las métricas — PT4. Tablas y gráficas de métodos y técnicas — PT5. Modelo(s) de regresión 	
<i>Productos de trabajo de salida</i>	
<ul style="list-style-type: none"> — PT5. Reporte de resultados del procesamiento de datos. 	

3.2.4 Proceso 4: Validación por estudio de caso

A través de este proceso, se proporciona una validación empírica y contextualizada, respaldada por la evidencia extraída de la implementación de estrategias de Social media de cada uno de los procesos anteriores. La Validación por Estudio de Caso no sólo examina la funcionalidad técnica de las herramientas y procesos, sino que también profundiza en su integración y adaptabilidad en escenarios dinámicos y cambiantes.

4. Validación	
<i>Propósito general</i>	
Verificar los procesos asociados a la recopilación y evaluación de datos durante el ciclo de vida del marco de trabajo.	
<i>Objetivos</i>	
O3.1 Evaluar la calidad de los experimentos de los procesos anteriores O3.2 Realizar el monitoreo y control de la red social	
<i>Roles</i>	<i>Herramientas sugeridas</i>
— Analista de redes sociales	<ul style="list-style-type: none"> — Hojas de cálculo — Entornos de desarrollo integrado para análisis estadístico de datos — Plataformas de software estadístico (IBM SPSS)
<i>Actividades</i>	
A4.1 Evaluar los procesos asociados al experimento y medir los resultados del experimento A4.2 Comparar las medidas contra otros experimentos A4.3 Optimizar el modelo para la mejora con monitoreo y control	
<i>Productos de trabajo de entrada</i>	
— PT7. Reporte de resultados del procesamiento de datos.	
<i>Productos de trabajo de salida</i>	
— PT8. Lecciones aprendidas	

Finalmente, este capítulo presenta el Marco de trabajo para el minado de Social media basado en GC (Figura 8), una estructura compuesta por los cuatro procesos que encapsula de manera integral las actividades fundamentales relacionadas con el conjunto de herramientas y los roles estratégicamente propuestos. Este marco de trabajo no solo representa un hito significativo

en la aplicación y comprensión de las técnicas y los algoritmos de Social media, sino que también proporciona una guía clara y detallada para los profesionales que buscan adentrarse en el análisis de datos en redes sociales.

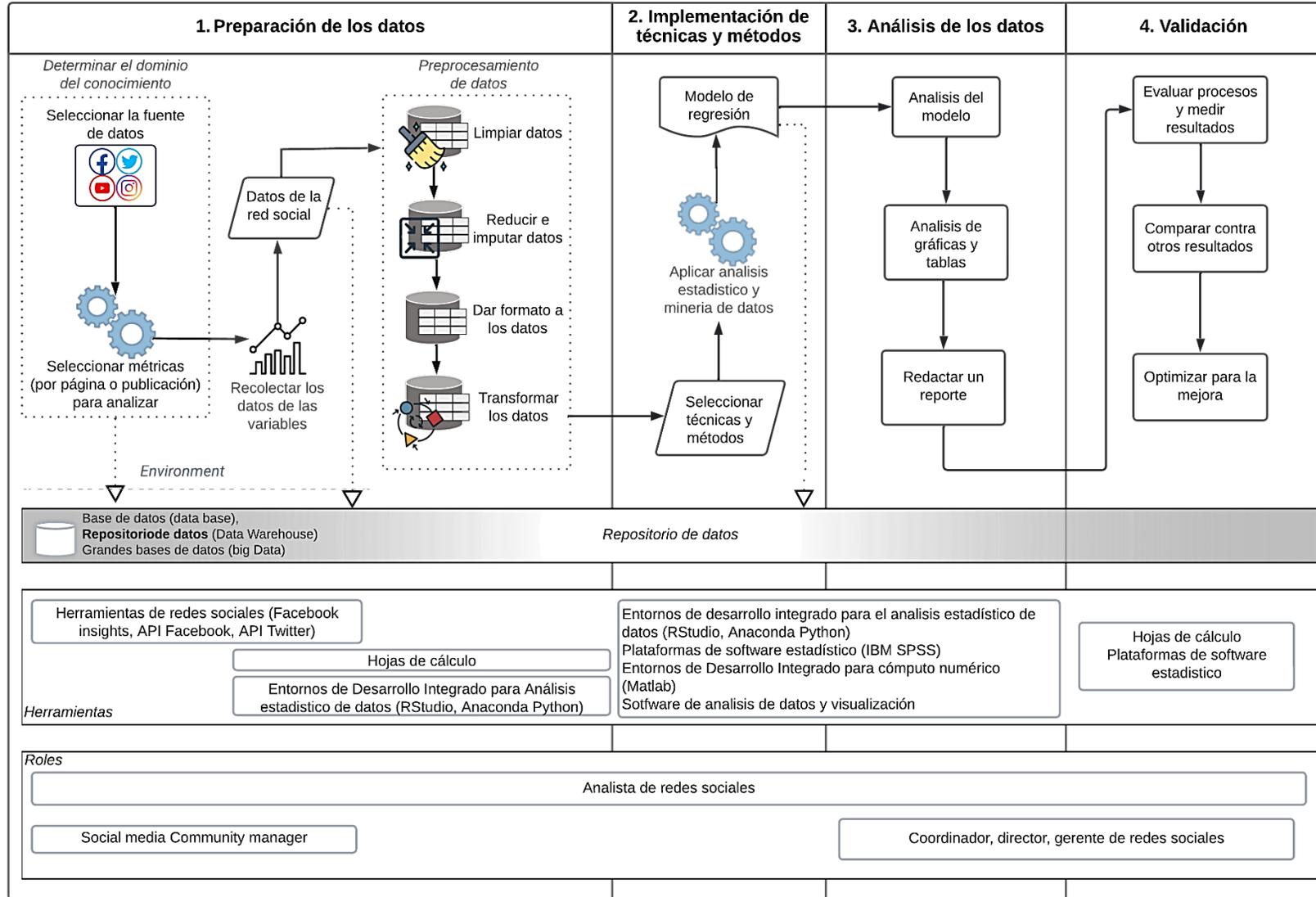


Figura 8 Marco de trabajo para el minado de Social media basado en GC.

CAPÍTULO IV: EXPERIMENTACIÓN

La aplicación del Marco de Trabajo presentado en el capítulo III en fan pages con características distintas, permitió realizar un análisis utilizando la importancia de las métricas extraídas y los cálculos de coeficientes de correlación de Spearman (Sharma & Lulandala, 2021). Los datos preprocesados de Facebook dieron como resultado propuestas para un modelo de regresión con un mejor ajuste para métricas como el User Engagement, los Clicks y el Alcance.

4.1 Fan page Mujeres en la Investigación - UABC (Parte 1)

En la fan page de Facebook para el programa de extensión de Mujeres en la Investigación - UABC (<https://www.facebook.com/mujeresinvestigacion>), desde el año 2013 se ha publicado información cultural, científica y tecnológica, en formatos de videos y fotografías de la realidad de trabajos de campo, información de tesis de posgrado, investigaciones, conferencias, pláticas, talleres, entre otros y compartido los mismos tipos de información de otras fuentes. El programa Mujeres en la Investigación ha contribuido activamente con el objetivo de generación, aplicación innovadora y transferencia del conocimiento y la promoción de la ciencia en diversos niveles educativos desde hace diez años. Por ello, se consideró de relevancia aprovechar estrategias para llegar a nuevos usuarios de la red social y obtener un mejor posicionamiento y visibilidad de los tipos de información compartida.

4.1.1 Análisis de las variables

En este primer acercamiento a los datos, se definió una estrategia para analizar datos dentro de un período de tres meses, todos ellos comprendidos en una serie de eventos de demostración y divulgación científica en formato online, organizados y publicados a través de esta fan page. Debido a las medidas de confinamiento por el Covid-19 y considerando el contexto educativo, se descartaron datos correspondientes al periodo marzo 2020 - enero 2022. Los datos fueron generados entre el 1 de febrero y el 30 de abril de 2022, correspondientes a 95 publicaciones. Teniendo en cuenta que la cantidad de personas alcanzadas por la publicación, los me gusta, los comentarios y la cantidad de veces que son compartidos cambia continuamente en las

publicaciones, se realizó una actualización de la muestra de datos al 30 de mayo de 2022 para tener el número más reciente de reacciones, comentarios, clics y compartidos en las mismas 95 publicaciones. Los autores Eriksson et al. (2019) y Laestadius (2017) recomiendan permitir un período para que los usuarios reflexionen sobre el contenido después de su publicación. Se decidió que un mes después de la última publicación debería ser suficiente, ya que la mayoría de las reacciones suelen recibirse dentro de las primeras 48 horas. El resultado de la extracción se concentró en un documento CSV. La tabla 13 presenta un ejemplo de los datos que se obtienen, donde cada tupla o renglón corresponde a una publicación realizada y las columnas a las diferentes variables que son proporcionadas por Facebook. Las variables pueden cambiar dependiendo de si los datos son extraídos desde una Fan page o un grupo de Facebook.

Tabla 13 Ejemplo de datos obtenidos desde Facebook.

Identificador	Título	Fecha publicación	Enlace	Tipo	Personas alcanzadas	Comentarios	Reacciones	Compartido
587269730	Actualización de estado	4/1/2022	https://www.facebook.com/mujeresinvestigacion	Foto	561	3	22	5
803869470	Actualización de estado	5/1/2022	https://www.facebook.com/mujeresinvestigacion	Foto	170	0	2	0
693718180	encuesta: fanpage mujeres en la investigación - uabc	7/1/2022	https://www.facebook.com/mujeresinvestigacion	Enlace	99	0	2	2
559681160	Actualización de estado	17/1/2022	https://www.facebook.com/mujeresinvestigacion	Foto	182	0	3	0
486311840	Actualización de estado	17/1/2022	https://www.facebook.com/mujeresinvestigacion	Foto	560	0	12	0
556192430	Fotos subidas con el celular	19/1/2022	https://www.facebook.com/mujeresinvestigacion	Foto	298	0	5	3
...

En el preprocesado de datos se utilizó la fórmula de Jayasigh (2019) para calcular la variable de salida User Engagement de las 95 publicaciones. La nueva variable calculada se utilizó

como un índice que midió el interés de los seguidores con las publicaciones. Posteriormente, se realizó un proceso de limpieza e imputación de datos faltantes siguiendo los criterios descritos en la Tabla 11, sustituyendo los datos faltantes por ceros. La Figura 9 muestra el algoritmo propuesto para obtener el resultado del User Engagement por publicación, donde para este primer experimento se eliminaron las publicaciones que tenían un User Engagement de cero, dejando 66 publicaciones en el resultado final. Al aplicar técnicas estadísticas, se realizó un análisis preliminar para verificar las distribuciones de normalidad y homogeneidad de varianza en la muestra con una prueba de Kolmogorov-Smirnov. Se utilizaron herramientas SPSS y Rstudio para apoyar el procesamiento de los resultados y la visualización gráfica.

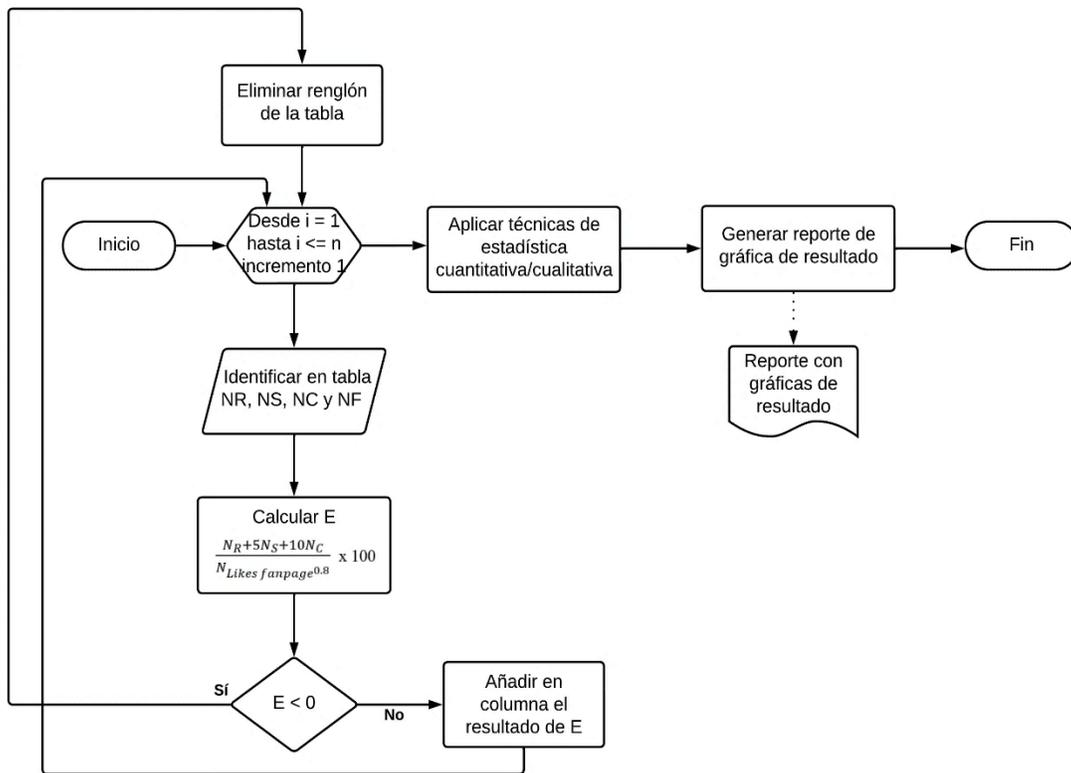


Figura 9 Diagrama de flujo de Engagement para fan page en Facebook.

De las variables, se seleccionaron tres según sus características: User Engagement (tipo continuo); Personas alcanzadas como un valor que incluye a todos los usuarios a los que les gusta la fan page, seguidores y otros usuarios que vieron las publicaciones (tipo cuantitativo y discreto); y Clics en la publicación para revisar con mayor detalle (tipo cuantitativo y discreto), dado que no

existe relación causal entre ellos. Las medidas de tendencia central permitieron resumir al conjunto de datos en un solo valor (Tabla 14).

Tabla 14 Diagrama de flujo de Engagement para fan page en Facebook.

Variable	Media	Mediana	Desviación estándar	Coefficiente de variación	Rango
User Engagement	66.2490	9.71	104.33	157.48%	0.32 – 506
Personas alcanzadas	970.326	355	1056.92	108.92%	97 – 4519
Clicks en la publicación	52.4631	8	76.31	145.46%	0 – 306

Los resultados de la media y mediana indican que los datos tienen un sesgo hacia valores cercanos a cero, por lo que la interacción de los usuarios en las publicaciones tiene un rango de 0.003 a 658 en promedio (Tabla 14). Esta información se ve confirmada por la alta desviación estándar y los coeficientes de variación, lo que representa una dispersión en el rango de los datos. La Figura 10 muestra el comportamiento de las variables en relación con la dispersión de los datos en la muestra. Se observa un comportamiento con tendencia lineal entre las variables al interactuar con los clics por publicación. Este comportamiento es más notorio entre el User Engagement y los clics en las publicaciones. Por lo tanto, se infiere que existe una relación entre la cantidad de reacciones, comentarios y la cantidad de veces que se comparte una publicación al momento de hacer clic para visualizarla. Esto significa que cuanto mayor sea el número de clics en una publicación, mayor será el valor de participación.

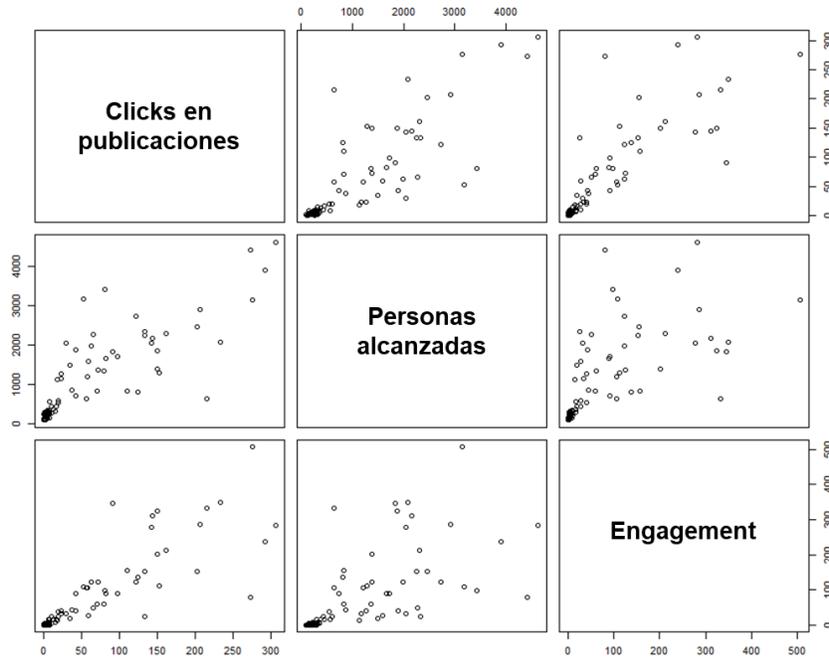


Figura 10 Diagrama de dispersión de las variables analizadas.

En el análisis cualitativo se consideraron los resultados de las medidas estadísticas exploratorias (Tabla 13) y Figura 10, se revisó la normalidad de los datos muestrales de las variables Engagement y Personas alcanzadas para determinar la tendencia que siguieron individualmente con el análisis de normalidad de los QQ-Graphs. La Figura 11 muestra las dos gráficas analizadas, donde ninguna de las variables sigue la línea recta, mostrando desvíos y colas en ambos extremos. Se nota un sesgo de los datos a la izquierda de la línea de tendencia, lo que indica que no hay normalidad en la distribución de los datos, lo que permite el uso de métodos de análisis no paramétricos.

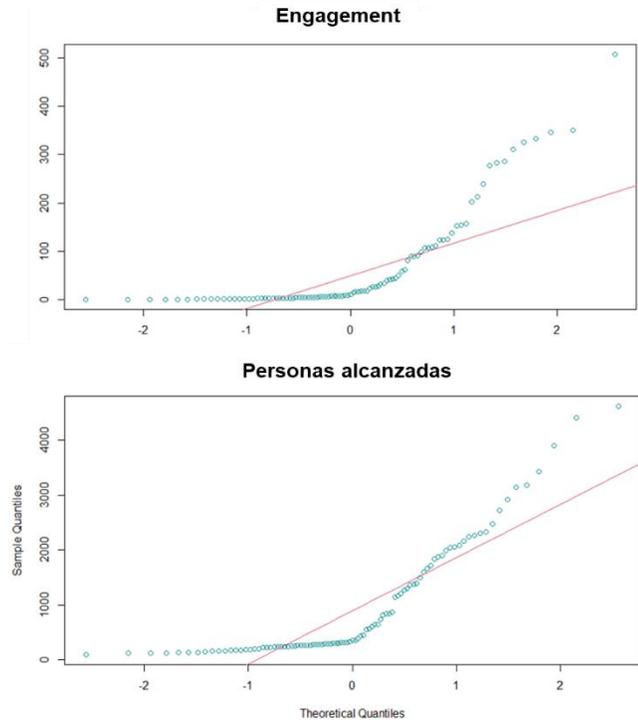


Figura 11 QQ-Graphs para verificación de normalidad de datos.

Para confirmar la no normalidad de la distribución de los datos y dada la cantidad de datos contenidos en la muestra, se utilizó Rstudio y la prueba de Lilliefors para contrastar si el conjunto de datos se ajustaba o no a una distribución normal (Gil-Quintana et al., 2022). Los resultados de la prueba de Lilliefors indicaron que para un nivel de significancia del 5% ($\alpha=0,05$), el p-valor calculado es menor ($2,2e-16$) en todas las variables analizadas; por lo tanto, los datos provienen de una distribución no normal.

De lo observado en el diagrama de dispersión (Figura 10), QQ-Graphs (Figura 11) y las estadísticas básicas (Tabla 14), se determinó que se realizaría una prueba de normalidad no paramétrica. Cuando la distribución de los datos no es normal, suelen ser análogos a las pruebas de tipo ANOVA. Por lo tanto, aplicar una prueba de χ^2 permitió, por contraste de hipótesis, probar afirmaciones sobre las funciones de probabilidad (o densidad) de las variables estudiadas (Al-Maatouk et al., 2020).

Se tomaron las variables categóricas Tipo de contenido y Origen del contenido publicado. Con estas columnas se realizó una categorización del 1 al 4 con los elementos descritos en la Tabla

15. Las categorías creadas corresponden a 1) fan page del caso de estudio donde se encuentra la creación de contenido original y de producción propia; 2) Contenido creado por la UABC; 3) Contenidos de otras universidades; y 4) Contenidos de otras páginas de fans de divulgación y divulgación científica.

Tabla 15 Variables de categorización de contenidos.

Tipo de contenido publicado		Origen del contenido publicado	
Elemento	Categoría	Elemento	Categoría
Foto	1	Estudio de caso	1
En vivo en Facebook	2	UABC	2
Enlace	3	Otras universidades	3
Vídeo	4	Otras fan pages	4

El resultado de aplicar estas dos categorizaciones originó la Tabla 16 de contingencia 4 con las condiciones para aplicar la prueba de χ^2 . El valor de chi-cuadrado calculado (12.434) con nivel de significancia ($\alpha = 0,05$) obtuvo un valor de p de 0,006, posicionando el chi-cuadrado calculado fuera de la distribución de la curva. Por lo tanto, se acepta H_a y existe relación entre ambos conjuntos con las variables categorizadas que fueron analizadas y son estadísticamente comparables. Por tanto, se valida una relación entre el tipo de elemento y el origen de lo publicado en la fan page.

Tabla 16 Tabla de contingencia para la prueba de normalidad chi-cuadrado.

Condición	Clase 1	Clase 2	Clase 3	Clase 4
1. Tipo de elemento	81	5	5	4
2. Origen del contenido	62	20	6	7

4.1.2 Modelo de regresión y Minería de datos

Del conjunto de datos obtenidos para el comportamiento de la página de fans (<https://data.world/paolavelazquez91/womens-in-research-uabc>), se seleccionaron las variables que no fueron utilizadas en el cálculo de User Engagement (reacciones, comentarios y compartidos) para evitar causalidad. Las variables de Clicks en publicación y Personas alcanzadas (Figura 10), se calculó el coeficiente de correlación de Spearman contra el User Engagement, encontrando que las variables analizadas mantuvieron una correlación positiva la cual se consideró

fuerte al ser superior a 0,8 en todas sus combinaciones. Se infiere que, a mayor número de clics, a más personas se llega y, por lo tanto, mayor número de reacciones, comentarios y compartidos, impactando directamente en el User Engagement.

Se realizó un análisis del modelo de regresión teniendo en cuenta las variables independientes disponibles, dando como resultado un modelo de regresión lineal múltiple para predecir el User Engagement que integra todo tipo de publicaciones a partir de las variables independientes Número de clics por publicación y Número de personas alcanzadas, y es definido de la siguiente manera:

User Engagement

$$= 10.61 + 1.34(\text{Clicks por publicación}) - 0.01(\text{Personas alcanzadas}) \quad (2)$$

Al calcular el valor de User Engagement, los rangos pueden variar desde 0% hasta al menos 100%. Se considera que una página de fans con una coherencia en rangos superiores al 70% de User Engagement en sus publicaciones tiene usuarios comprometidos (Shahbaznezhad et al., 2021). El modelo de regresión definido (ecuación 2) obtuvo un coeficiente de determinación múltiple (R^2) de 0,79 de influencia de las variables independientes (p-valor de 2,2e-16). La ecuación muestra que el Engagement está muy influenciado por la interacción, mediante clics, con una publicación. También indica que el Engagement disminuye por cada persona alcanzada por la publicación que no interactúa con la publicación. Se seleccionó uno de los posts y mediante el modelo de regresión se determinó que el Engagement en una publicación con 3249 personas alcanzadas y 207 clics tendría un valor de 255,53%. Esto se considera altamente representativo y permite predecir el Engagement en publicaciones futuras en función de los Clics por publicación y las Personas alcanzadas.

Por otro lado, en la Figura 12 se muestra qué tipo de contenido se compartió en la fan page según las categorías asignadas (Tabla 14). Las burbujas azules corresponden al contenido que es completamente de producción propia, creado a partir de actividades realizadas por la fan page; esto significa que los usuarios comparten y consumen principalmente la información que ellos mismos producen en la página de fans.

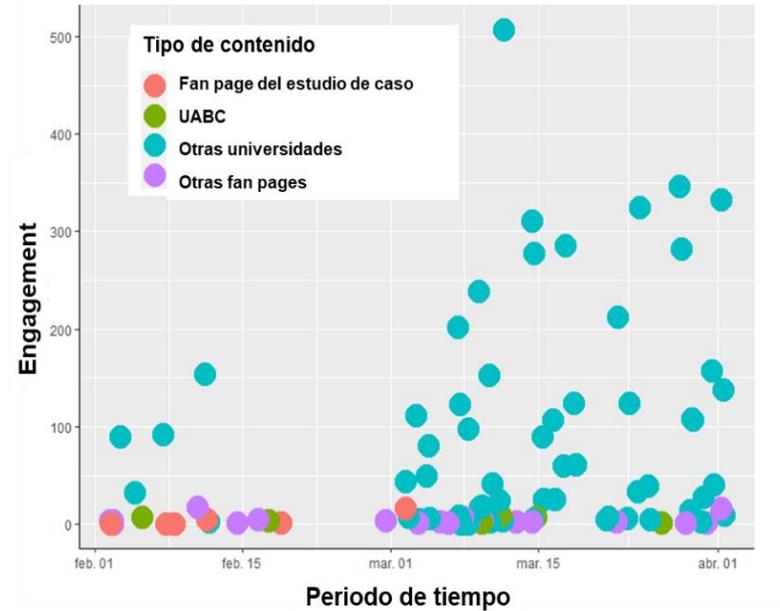


Figura 12 Gráfico de burbujas con los tipos de contenidos que comparte la fan page.

El análisis de minería de datos permite descubrir características ocultas en el conjunto de datos. La Figura 13 muestra el gráfico bivariado obtenido de la aplicación de k-medias, donde se visualizan los clusters para todas las variables obtenidas en la minería de datos de Facebook representando las reacciones, interacciones, comentarios, personas alcanzadas, tiempos compartidos y clics por publicación. Todos los datos están representados por puntos en el gráfico, utilizando dos componentes principales. Los símbolos de color verde indican a qué grupos pertenecen. Estos puntos de datos no son los datos originales extraídos de las variables, pero cada uno representa un resumen de los datos subyacentes. En este caso se ha explicado el 92.17% de la variabilidad de puntos dentro de la muestra de datos.

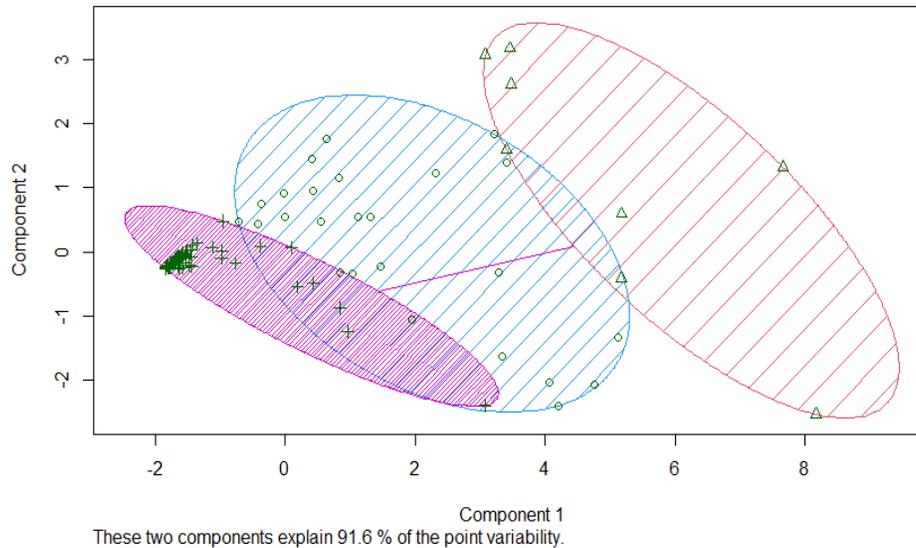


Figura 13 Gráfica de variables extraídas de Facebook con la técnica de k-medias para k=3.

4.2 Fan page Mujeres en la Investigación - UABC (Parte 2)

La estrategia definida para un período de dos años fue que la selección de los datos extraídos fuera de una serie de eventos de demostración y divulgación científica en formato online, organizados y publicados a través de esta fan page. Los datos fueron generados entre el 1 de enero de 2021 y el 31 de diciembre de 2022, correspondientes a 345 publicaciones. Teniendo en cuenta que la cantidad de me gusta, comentarios y compartidos cambia continuamente en las publicaciones, se realizó una actualización de la muestra de datos al 30 de mayo de 2023 para tener el número más reciente de reacciones, comentarios, clics y compartidos en las mismas 345 publicaciones. Los autores Agusriandi et al., (2020) y Niciporuc (2014) recomiendan permitir un período para que los usuarios reflexionen sobre el contenido después de su publicación. Se decide que un mes después de la última publicación debería ser suficiente, ya que la mayoría de las reacciones suelen recibirse dentro de las primeras 48 horas. El resultado de esta fase generó un documento XLSX o CSV donde cada tupla corresponde a una publicación y las columnas contienen la información de fecha, tipo de publicación, clics, reacciones, comentarios y compartidos.

4.2.1 Análisis de las variables

El cálculo del User Engagement de cada publicación se realizó mediante la fórmula propuesta por Jayasingh. Los resultados de User Engagement se colocaron como una nueva variable para analizar el índice de User Engagement de los seguidores hacia la fan page. Posteriormente se realizaron los procesos de limpieza e imputación de datos faltantes haciendo uso de la tercera forma, donde con el cálculo del promedio de los valores existentes se colocaron los datos faltantes. Al aplicar técnicas estadísticas, se realizó un análisis preliminar para verificar las distribuciones de normalidad y homogeneidad de varianza en la muestra con una prueba de Kolmogorov-Smirnov (Oviedo-García et al., 2014; Bonsón & Ratkai, 2013). Se utilizaron las herramientas SPSS y RStudio para apoyar el procesamiento de resultados y la visualización gráfica.

La muestra de exploración inicial incluyó la caracterización del conjunto de seguidores de la fan page. Aunque el contenido generado en dicha página es de mujeres. La Figura 14 presenta los 1,532 seguidores, el 90% pertenece a localidades de México, donde el 83% corresponde a mujeres en el rango de edad entre 25 - 50 años y el 17% de seguidores masculinos distribuidos entre 25 - 50 años. 45 años en su mayoría. La distribución de seguidores es comprensible debido al objetivo de la fan page que es posicionar la difusión de contenidos científicos creados por mujeres investigadoras.

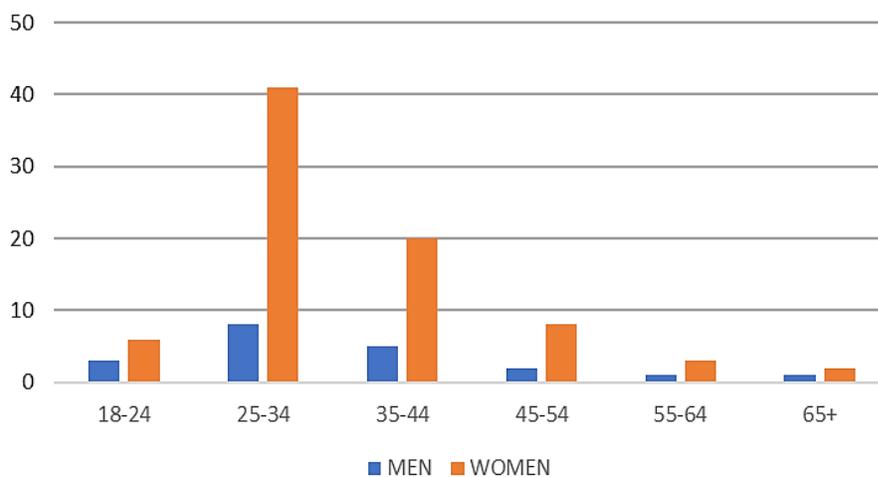


Figura 14 Distribución por edades de los seguidores de las páginas de fans.

Durante el período de 24 meses comprendido entre enero de 2021 y diciembre de 2022, la muestra registró un total de 345 publicaciones. La Tabla 17 muestra la distribución de tres métricas cuantitativas seleccionadas y extraídas del conjunto de datos: número de comentarios, número de reacciones y número de veces que se compartió una publicación, separados por meses de cada año. Centrándonos en las métricas de participación de todos los usuarios, hubo un aumento.

Tabla 17 Distribución de reacciones, comentarios y acciones extraídas del conjunto de datos.

2021				2022			
Mes	Métricas			Mes	Métricas		
	Reacciones	Comentarios	Compartidos		Reacciones	Comentarios	Compartidos
1	29	0	11	1	771	165	59
2	61	2	17	2	472	56	57
3	79	2	12	3	5996	994	299
4	254	25	52	4	393	507	71
5	48	1	2	5	122	181	11
6	7	0	1	6	373	415	4
7	252	103	23	7	30	1	2
8	0	0	0	8	43	1	0
9	40	1	0	9	22	1	5
10	7	1	0	10	11	2	0
11	83	16	2	11	7	0	2
12	8	1	0	12	19	1	1
TOTAL	868	152	120	TOTAL	8259	2324	511

Esto significa que las reacciones aumentaron un 951,5%, los comentarios un 1528,9% y las acciones un 425,8%. El tercer mes (marzo de 2022) mostró la mayor interactividad. Esto se debió a la realización de un evento en streaming en vivo para difusión de la ciencia con 7289 interacciones en publicaciones, a diferencia de marzo de 2021 con 93 interacciones. Entre los hallazgos más interesantes está que en 2021 el séptimo mes fue el que tuvo más interacciones con una sola publicación realizada. El mes 4 en ambos años mantuvo una interactividad similar, ya que son actividades las que realiza la fan page todos los años. En la mayoría de los meses se observa aumento, pero hay casos particulares donde se observó una disminución, como los meses 9, 10 y 11 del año 2022.

Se identificaron cinco categorías diferentes de publicaciones: 1) Video, 2) Foto, 3) Enlace, 4) Video en vivo y 5) Texto. En la Tabla 18 se describe la distribución de puestos según las categorías creadas. Al profundizar en las diferentes categorías, surgen hallazgos significativos: en 2021, el mes con mayor número de interacciones (Tabla 16) fue julio, pero solo tuvo una publicación tipo Foto. En el caso del año 2022, la concentración de más interacciones fue en marzo,

donde cerca del 85% de las publicaciones fueron elementos de la categoría fotografía. Se identificó que el número de publicaciones de tipo Texto disminuyó considerablemente. Las publicaciones de tipo vídeo también disminuyeron, pero aumentaron los enlaces a otros elementos externos. Las publicaciones con fotos se mantuvieron constantes. Se separa por mes con el fin de visualizar el comportamiento de los conjuntos de categorías de forma ordenada.

Tabla 18 Porcentaje de publicaciones en las categorías.

Mes	2021					Mes	2022				
	Categorías						Categorías				
	1	2	3	4	5		1	2	3	4	5
1	6.66	7.76	17.65	4.17	0	1	0	14.73	15.79	0	0
2	13.33	24.14	29.41	29.17	16.67	2	0	11.63	15.79	25.00	0
3	6.66	12.93	23.53	8.33	16.67	3	40	48.84	15.79	12.50	25
4	46.66	17.24	11.76	29.17	0	4	0	7.75	10.53	25.00	25
5	13.33	12.07	5.88	16.67	0	5	20	5.43	5.26	37.50	25
6	6.66	1.72	5.88	0	0	6	20	3.88	5.26	0	0
7	0	0.86	0	0	0	7	0	0.78	10.53	0	0
8	0	0	0	0	16.67	8	0	1.55	0	0	0
9	0	8.62	5.88	4.17	0	9	0	1.55	5.26	0	0
10	6.66	4.31	0	4.17	0	10	0	0.78	10.53	0	25
11	0	7.76	0	4.17	33.33	11	0	1.55	0	0	0
12	0	2.59	0	0	16.67	12	20	1.55	5.26	0	0

En la Tabla 19 se comparan los resultados de ambos años separados por meses. Se puede observar la diferencia entre los valores, destacándose los primeros seis meses del año 2022. Se calculó con ambas fórmulas en los dos períodos y se compararon entre ellos. En general, si hay un aumento de un año a otro, independientemente de la fórmula. En el mes once hay una disminución solo en la participación en Facebook. A medida que hay más interacciones, se genera más User Engagement. Es posible observar que en 2021 el Engagement con la fórmula, calculado por Facebook, es superior al obtenido con las publicaciones de 2022. En el mes 8 de 2021 las publicaciones no tuvieron interacciones impactando el User Engagement. Mientras que el mes 7 obtuvo el valor más alto en este periodo. La Tabla 20 nos permitió observar las correlaciones entre las diferentes categorías extraídas del conjunto de datos. Esto ayudó a verificar la relación entre tipo de contenido y período de tiempo como variables relacionadas, donde la incidencia es visible en los elementos de tipo foto. Además, en la Tabla 3 podemos ver que en el cálculo del User Engagement con fórmula propuesta por Jayasigh (2019) y Facebook. Los contenidos de tipo Texto serían los que menos influencia tienen; mientras que los Contenidos de Vídeo y Enlaces generan un impacto medio en la relación.

Tabla 19 Comparación de los dos años entre las fórmulas por mes.

Mes	Engagement 2021		Engagement 2022	
	Facebook	Jayasigh	Facebook	Jayasigh
1	2.8571	2.9460	45.2272	29.4831
2	2.3488	1.8079	29.25	19.6131
3	3.2608	2.4640	102.662	58.3242
4	9.1944	7.4097	64.7333	71.9644
5	2.4285	1.0314	24.1538	25.9515
6	2	1.2610	113.1429	105.4626
7	378	295.8294	11	5.4398
8	0	0	22	7.1212
9	3.4166	1.1126	9.3333	7.6158
10	1.1428	0.5086	3.25	1.5577
11	6.3	3.1155	4.5	4.0057
12	2.25	0.9643	5.25	2.5221

Tabla 20 Comparación de correlaciones de Spearman por categoría.

Fórmulas Engagement	User	Categorías				
		1) Vídeo	2) Fotos	3) Enlace	4) Vídeo en vivo	5) Texto
Facebook		0.41	0.78	0.42	0.47	0.16
Jayasigh		0.33	0.77	0.32	0.48	0.21

4.2.2 Modelo de regresión y Minería de datos

Al realizar un cluster con un total de publicaciones que se realizaron distribuidas mediante los distintos cálculos de las ecuaciones implementadas. La Figura 15 distribuye las publicaciones de menor a mayor participación del usuario con ambas fórmulas. Por el método del codo se reducirá que el número adecuado de grupos sea tres. El primer grupo (azul) corresponde a publicaciones que son bajas debido a la interactividad y por consecuencia tienen poca participación del usuario. El grupo dos (verde) tiene las publicaciones con un compromiso entre 100 y 300. Finalmente, el grupo tres (rojo) representa las publicaciones con un compromiso mayor a 300 por tener mayor interactividad. Se puede observar en el gráfico que existe una mayor concentración de publicaciones en el grupo uno, denotando que la mayor parte del compromiso en las publicaciones no tiene la respuesta esperada. Es importante mencionar que el grupo dos contiene las publicaciones correspondientes a un evento de divulgación científica. En un análisis más detallado de las publicaciones, la Figura 16 establece que los clusters identificados son los mismos

que los obtenidos en el análisis anterior; es decir, el grupo 1 del gráfico de la Figura 15 es el mismo grupo 1 de la Figura 16.

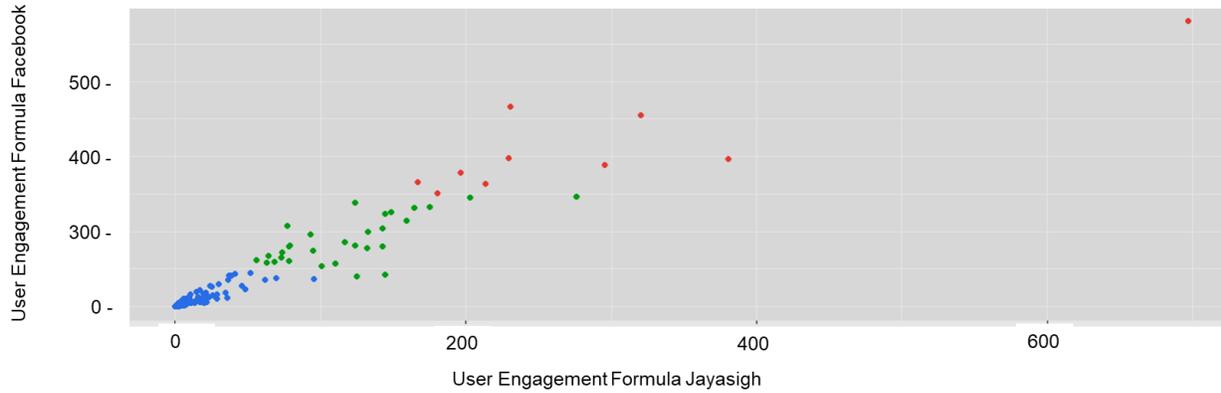


Figura 15 User Engagement con las dos fórmulas.

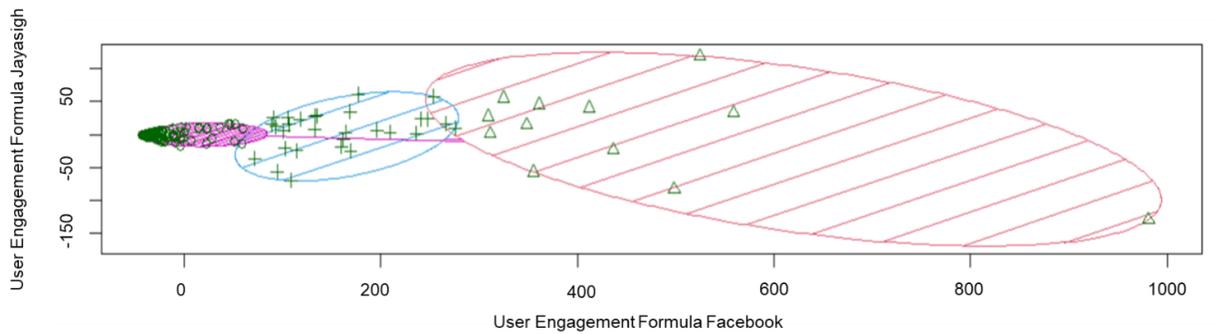


Figura 16 Agrupamiento de User Engagement con las dos fórmulas.

El primer grupo explica los elementos que la fórmula de Jayasigh (2019) y la proporcionada por Facebook tienen baja interacción y por tanto poca participación del usuario. En el segundo, tenemos un User Engagement de 100 a 300 de Facebook, pero con la propuesta de Jayasigh en 50. Del conjunto de datos obtenidos para el comportamiento de la fan page, se tomaron las variables que se utilizaron para clasificar por tipo y número de interacciones. Se seleccionó el cálculo de User Engagement. Por lo tanto, se infiere que a mayores interacciones y depende del tipo de contenido, el valor de User Engagement se ve afectado.

Se extraen análisis cualitativos aplicados en datos no estructurados, relaciones y grupos que permitieron aprovechar no sólo las descripciones obtenidas sino también las relaciones y reglas. De esta manera, el análisis cualitativo apoya la identificación de varios tipos de relaciones dentro de la interactividad de los datos en la red social Facebook. Se realizó un análisis de modelo de regresión teniendo en cuenta la cantidad de interacciones en cada publicación y el tipo de contenido (foto, vídeo, texto, enlace) como variables independientes por tener una mayor correlación que no expresaba causalidad para la variable de User Engagement, dando como resultado un modelo de regresión lineal múltiple para predecir la participación del usuario, y se define de la siguiente manera:

$$User\ Engagement = 0.56_{Interactions} - 0.29_{Photo} - 1.58_{Text} + 172.18_{Video} + 22.23_{Live} \quad (3)$$

El modelo de regresión definido (ecuación 3) obtuvo un coeficiente de determinación múltiple (R^2) de 0,59 de influencia de las variables independientes (p-valor de 2,2e-16). La ecuación muestra que la participación del usuario está muy influenciada por el tipo de contenido del video, sin importar si es en vivo o no. También indica que la participación del usuario disminuye con cada tipo de contenido de foto o texto que tiende. Esto se considera altamente representativo y permite interpretar la participación del usuario en futuras publicaciones en función del tipo de contenido y la cantidad de interacciones. Al calcular el valor del User Engagement, se considera que una página de fans con consistencia en rangos de User Engagement en sus publicaciones tiene usuarios engagement (Ballesteros Herencia, 2018). El modelo de regresión permitió interpretar el impacto de las categorías por tipo de contenido. Esto se considera altamente representativo y permite predecir la participación del usuario en publicaciones futuras según el tipo de contenido.

4.3 Cultura UABC

Se obtuvo una muestra no probabilística de 39 meses por conveniencia del periodo de marzo de 2020 a junio 2023 con un total de 11,507 publicaciones, de las fan page: 1) Cultura UABC, 2) Editorial UABC, 3) Feria Internacional Del Libro UABC y 4) UABC Radio. El programa de extensión de la cultura y la Ciencia, tiene como objetivo incentivar la cultura y divulgar el conocimiento técnico científico, definida en la agenda institucional 2023 – 2027.

4.3.1 Análisis de las variables

En la tabla 21 es posible observar el total de publicaciones obtenidas de cada página en el periodo asignado para un total de 11,507 publicaciones.

Tabla 21 Descripción de muestra del programa de extensión de la cultura y la Ciencia.

Página	Seguidores	Enlace	Periodo	Total Publicaciones
Cultura UABC	18 mil	https://www.facebook.com/CulturaUABC/	Marzo 2020 Junio 2023	1256
Editorial UABC	2 mil	https://www.facebook.com/editorialuabc	Marzo 2020 Junio 2023	427
Feria Internacional Del Libro UABC	27 mil	https://www.facebook.com/filuabcml	Marzo 2020 Junio 2023	8836
UABC Radio	22 mil	https://www.facebook.com/uabcradio	Marzo 2020 Junio 2023	988
			TOTAL	11,507

Los valores descriptivos de las variables del conjunto de las 11507 publicaciones se encuentran con diversos rangos, presentados en la Tabla 22. Los cuartiles dejan incidencia de que existe un sesgo en los datos inclinado hacia los valores menores a 15, dado que el primer cuartil donde se explica el 25% de los datos como menor que o igual a cero en caso de comentarios y compartidos, y los terceros cuartiles con el 75% de los datos es menor que o igual a este valor a 15 en el caso de las reacciones. Por otro lado, los valores máximos, se consideran valores atípicos y se analizarían por separado considerándolos como publicaciones con mayor éxito. Dado que las medias y medianas se mantienen cercanas a cero.

Tabla 22 Estadísticos básicos.

Variable	Valor min	Valor max	mediana	media	1er cuartil	3er cuartil
Reacciones	0	2503	5	25.46	2	15
Comentarios	0	955	0	9.96	0	2
Compartidos	0	1792	1	14.87	0	6

El histograma generado con el cálculo del Engagement permitió constatar que los datos si tienen un sesgo marcado hacia la izquierda, clara incidencia en publicaciones con menos de 200 en el valor del Engagement. La muestra es de 3569 después de eliminar las publicaciones repetidas al aplicar el algoritmo de la Figura 9, y 2932 de estas publicaciones tienen el comportamiento explicado.

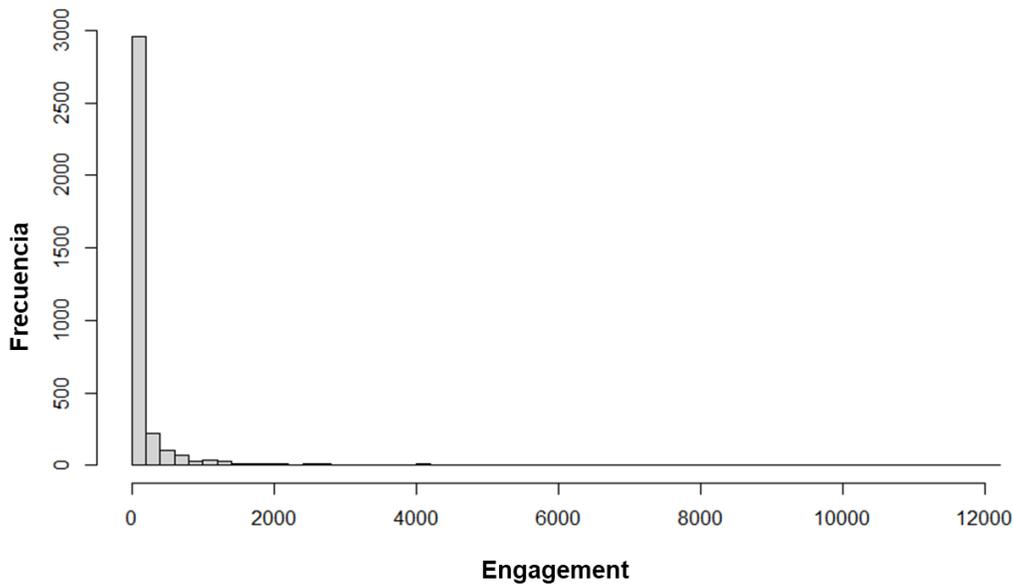


Figura 17 Histograma de Engagement

4.3.2 Análisis de la comunidad

La detección de comunidades en redes sociales plantea un desafío en la evaluación de la calidad de la estructura de la comunidad en la red social, debido a los enfoques para evaluar un esquema de partición al tratar los grafos que involucran sólo un tipo de actor (Dey & Roy, 2016). En este sentido, el análisis de redes sociales se aplica a diversos enfoques, los problemas con cualquier investigación pueden verse influenciados por el tamaño de la interacción entre la relación del actor (Biswal & Muley, 2018). A partir de la fan page de cultura UABC, se identificó la información de 30 fan page de Facebook en donde se promueven la cultura y divulgación de la ciencia ante la sociedad, las cuales permiten generar un conjunto de grafos e identificar cómo es la interacción entre ellas. Los algoritmos Fruchterman-Reingold, Yifan Hu y Noverlap (Tabla 1) se aplicaron en el análisis de si se presenta una relación o interacción entre una fan page como nodo central y otras fan page como miembros de una misma comunidad universitaria.

Se implementó una metodología adaptada de los autores Diliana & Indrawati (2022); Wahyuningtyas et al. (2021); Bedi & Sharma (2016); Gupta & Gupta (2016) y Kuz et al. (2016), considerando las actividades para Facebook de Velazquez-Solis et al. (2022) que se divide en cuatro fases: 1) Selección, 2) Construcción, 3) Modelado y, por último, 4) Análisis. El flujo de trabajo del método se muestra en la Figura 18.

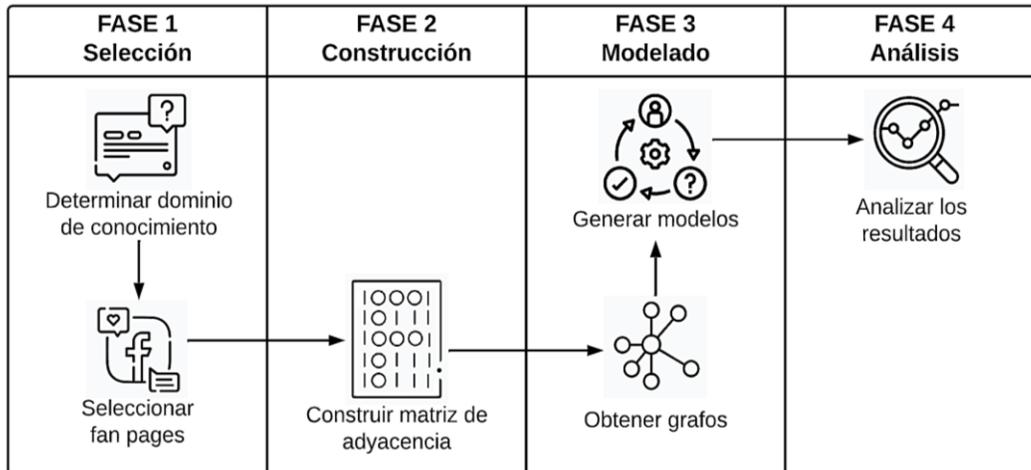
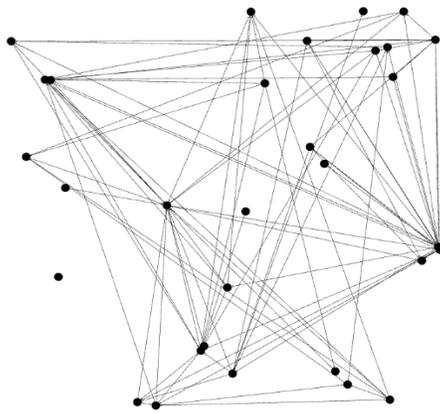
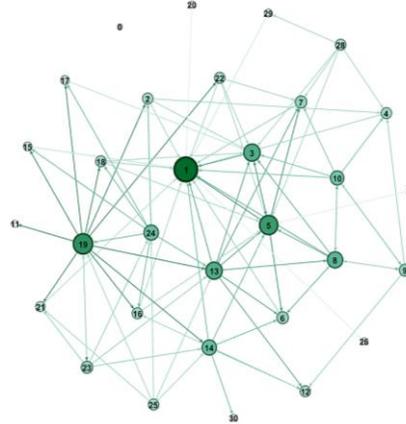


Figura 18 Metodología para la identificación de comunidades en Facebook.

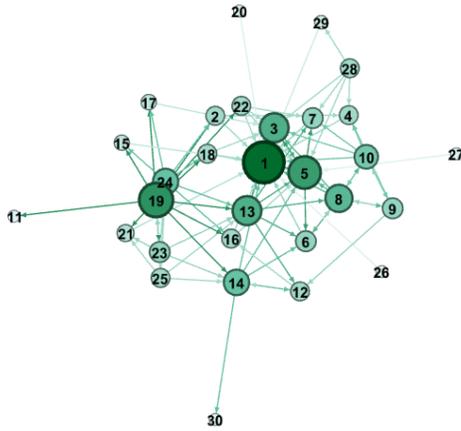
Donde primeramente se determinó que el área de especialización relacionada con el entorno de la aplicación de la información (Velazquez-Solis & Flores-Rios, 2021), como dominio del conocimiento, proporcione las características iniciales para la selección de las 30 fan pages oficiales de Facebook en una Institución de Educación Superior pública relacionadas con las actividades de promoción de la cultura y divulgación de la ciencia desde hace dos años. Posteriormente, para construir la matriz de adyacencia se inició con el análisis del conjunto de datos obtenidos en la fase anterior se mejoró con la visualización de la red utilizando la herramienta Gephi versión 0.10.1, para ello primero fue necesario adaptar los datos en R a un formato importable (Hanneman & Riddle, 2009). En Gephi se requiere de exportar dos archivos en formato CVS, uno con las 30 fan pages (nodos) y otro con la matriz de adyacencia que representaba las direcciones de las aristas del grafo como relaciones entre los nodos. Una vez ingresados los archivos, se configuran los parámetros de construcción para el grafo. Con la matriz de adyacencia y mediante el uso de tres algoritmos se obtuvieron un conjunto de grafos formados por 30 nodos y las 96 aristas iniciales que contenía la interacción de la comunidad universitaria. Se aplicaron los algoritmos de distribución Fruchterman-Reingold, Yifan Hu y Noverlap para tener una visualización más clara de la red y apreciar los roles que juegan los diferentes nodos. La Figura 19 es una representación de las interacciones entre las 30 fan pages de Facebook, haciendo que en cada fan page se represente la relación del seguidor que hace mención a otra fan page.



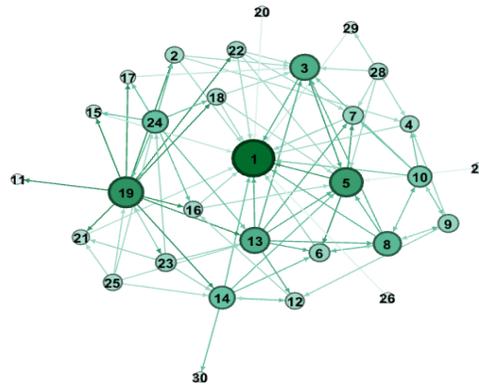
a) Grafo generado por default



b) Grafo generado con el algoritmo Fruchterman-Reingold



c) Grafo generado por el algoritmo Yifan hu



d) Grafo generado por el algoritmo Noverlap

Figura 19 Grafos generados por Gephi.

En la Figura 19 (a) se observa el resultado por default al cargar los datos, proporcionando únicamente la visualización de las relaciones con las direcciones de las aristas. En este caso las relaciones son visibles, pero algunos nodos están ocultos detrás de otros y existe incertidumbre sobre cuáles nodos tienen mayor fuerza. En la Figura 19 (b) se aprecia el resultado obtenido al utilizar el algoritmo Fruchterman-Reingold, el cual trata a las conexiones del grafo de manera similar, atrayendo a los nodos que están conectados entre sí y, simultáneamente, aplicando una fuerza de repulsión a todos los nodos para mantenerlos separados. Como resultado típico, este algoritmo genera un grafo con conexiones de tamaño similar y dispone los nodos en una

disposición circular, donde aquellos con más conexiones ocupan el centro y los nodos con menos conexiones se sitúan en las periferias. Al aplicar este algoritmo se puede observar cómo el nodo 1 – Cultura UABC, con mayor tamaño, no se encuentra céntrico, dado que las aristas en su mayoría se dirigen hacia él. El caso contrario en el nodo 13 – Vicerrectoría UABC Mexicali, donde el tamaño es asignado por las relaciones que éste mismo genera al seguir al resto de las fan pages. De esta misma forma, el nodo 19 - Centro Deportivo UABC es expulsado hacia el lado izquierdo por la fuerza de las aristas, porque a pesar de tener diversas conexiones son relaciones que este mismo ha buscado, sin que los otros nodos se relacionan con él.

En la Figura 19 (c), el algoritmo Yifan hu, muestra otro tipo de relación entre las fan pages donde los nodos se aprecian más oscuros entre mayor sea la cantidad de aristas, puesto que se trata de un algoritmo de repulsión. Es más visible que la fan page que no sigue a ninguna se aísla en un extremo, mientras que aquellas que tienen mayor cantidad de seguidores entre la lista de nodos, tienden a aproximarse al centro y crecer en diámetro, tal es el caso del nodo 1 – Cultura UABC, 3 – Extensión Cultural UABC Tecate y 13 – Vicerrectoría UABC Mexicali. El algoritmo Noverlap, presentado en la Figura 19 (d) permite ajustar el radio de separación entre nodos y la velocidad de ejecución. Al ser un algoritmo que evita que no exista una superposición entre los nodos, disminuye la posibilidad de que los nodos de un tamaño grande oculten a nodos pequeños (Humberstone, 2019). A diferencia de los algoritmos anteriores, para usarlo se debe de aplicar antes otro algoritmo que determine la configuración de partida, en este caso los algoritmos Fruchterman-Reingold y Yifan hu. Si el algoritmo Noverlap se ejecuta sin aplicar otro primero, se obtiene un resultado más alejado de esta configuración de partida (Morales-i-Gras, 2021). Esto puede implicar tiempo de computación muy elevado, sobre todo en redes masivas, lo cual en este caso no se presentó, y permitió observar cómo los nodos que tienen más relaciones crecen, nodo 1 - Cultura UABC, nodo 5 – Vicerrectoría Campus Tijuana, nodo 13 – Vicerrectoría UABC Mexicali y nodo 19 - Centro Deportivo UABC; mientras que se separan de los que tienen menor cantidad de aristas nodo 0 - UABC, nodo 27 – Sorteos UABC y nodo 30 - Instituto de Ingeniería UABC. Se esperaba que el nodo 0 fuera el centro de la comunidad, debido a las características del tipo de información institucional que se presentan en las publicaciones, pero en los tres algoritmos implementados, se mantuvo aislado. El nodo 23 correspondiente a las actividades de la Feria Internacional del Libro es el nodo con más popularidad entre los usuarios internos y externos a la

universidad, pero dentro de la comunidad analizada no cuenta con la suficiente fuerza en sus aristas para mantenerse centralizado.

Finalmente, como parte de los resultados se aplicó un ajuste de etiquetas para la generación de una nube de palabras con los 30 nodos (Figura 3), nuevamente se tomó como referencia el total de 96 aristas y la cantidad de aristas que recibía cada nodo para generar la comunidad. Sobresalieron los nodos referentes a 1 - Cultura UABC, 3 - Extensión cultural Tecate y 5 - Vicerrectoría Tijuana, debido a que éstos nodos cuentan con una mayor fuerza de atracción que el resto.



Figura 20 Cloud Word de la comunidad en Facebook para la Cultura y Divulgación de la Ciencia.

En el caso particular del conjunto de datos analizados, por tratarse del análisis de solo una comunidad, la información proporcionada inicialmente es confirmada y extendida en la visualización de los grafos, obteniendo un modelo de la comunidad que describe cuales fan pages tienen más relevancia y comunicación entre la sociedad. En el caso que el conjunto de datos sean perfiles de usuario públicos o grupos de usuarios el tratamiento de la información dependerá del dominio del conocimiento y las adecuaciones a la metodología.

CAPÍTULO V: CONCLUSIONES, APORTACIONES Y TRABAJO FUTURO

5.1 Conclusiones

La comunidad científica considera que el análisis de datos de las redes sociales es un área de investigación multidisciplinar que permite la creación, extensión y adaptación de métodos (Savelev et al., 2021) y modelos de análisis de datos que empíricamente proporcionan resultados cuantitativos utilizando métodos estadísticos (Velazquez-Solis et al., 2022). El implementar métodos de análisis estadístico y análisis cualitativo para establecer un marco de trabajo que permita entre otras cosas el cálculo de User Engagement o la interactividad en una fan page de Facebook para contenidos de divulgación científica permitirá trazar estrategias en función a los hallazgos. Se demostró que la cantidad de clics es la métrica más efectiva en comparación con las métricas habituales de interactividad, como la cantidad de comentarios, la cantidad de reacciones y la cantidad de veces que se comparte una publicación.

Para los roles involucrados en la gestión de redes sociales (Community manager o Social media manager), los resultados de esta investigación pueden servir como una idea de cómo aumentar el nivel de User Engagement en las publicaciones entendiendo a los usuarios interesados en temas científicos. El resultado del modelo de regresión se puede utilizar como método para seleccionar las estructuras de puestos más fructíferas o atractivas. Junto con otras métricas de Facebook, como el crecimiento de seguidores y los nuevos Me gusta, ayudan a comprender la participación del usuario y agregan precisión al predecir las necesidades específicas de cada fan page en el futuro.

El objetivo fue analizar la participación de los seguidores y el User Engagement en el contenido de los posts de una fan page con temas de divulgación científica. La aportación de esta investigación logró generar estrategias de divulgación científica para aumentar la visibilidad del contenido científico y mejorar el interés de la comunidad educativa, aprovechando la versatilidad de Facebook. Para este trabajo se destacan las siguientes conclusiones:

(1) La investigación de las diferentes fórmulas de User Engagement para obtener métricas es de gran valor porque incluye la visualización del contenido y la interactividad que tienen los usuarios en la fan page de Facebook.

(2) Existe una correlación positiva entre el User Engagement que recibe una publicación y el tipo de contenido al que se asocia, en función al tipo de propósito de publicación y el formato que esta tenga.

(3) En el análisis y visualización de las variables de cada publicación también se aplicó el clustering como técnica de minería de datos, demostrando que el contenido de producción propia proporciona un mayor User Engagement y, por ende, una mayor interactividad y posibilidad de viralización.

(4) Se propone el uso de modelos de regresión para establecer una interpretación o predicción de las métricas obtenidas de Facebook. La creación de nuevos modelos para la interpretación de las variables de Facebook dio como resultado dos modelos de regresión lineal múltiple que permiten identificar el User Engagement y el alcance en función de las diferentes categorías en cada publicación.

(5) El Marco de trabajo para el minado de Social media basado en GC permite dar seguimiento desde un conocimiento de novato hasta experto. Proporcionando una guía para la aplicación de técnicas de Social media.

En consecuencia, es recomendable utilizar el análisis estadístico cuantitativo y cualitativo de datos de Facebook, metodologías de análisis de datos de redes sociales y un modelo de regresión lineal como herramientas estratégicas para mejorar las prácticas de creación y difusión de contenidos enfocados en temas de Ciencias e Ingeniería. Para esta investigación se destaca la importancia de un adecuado tratamiento y análisis de los datos de las redes sociales, determinado por el tipo de seguidores y consumo de las publicaciones creadas.

5.2 Aportaciones

Las principales aportaciones de esta investigación se enumeran a continuación:

- Se presenta una comprensión general de Social media a partir de definir la diversidad de características, técnicas de aplicación y los procesos que involucra. Permitiendo la concepción de una idea concisa sobre el Social media en el ámbito de las redes sociales, así como los elementos centrales entorno.
- A través de la aplicación de la metodología de intervención se presenta un enfoque diferente para el conocimiento de lo previamente identificado en el marco de trabajo, donde el objetivo permitió identificar los roles, actividades, productos de trabajo y herramientas tecnológicas disponibles para los procesos de Social media. Se presenta una metodología detallada, con un nivel de abstracción que permite su posterior aplicación en estudios de otras disciplinas.
- La definición del Marco de trabajo para el minado de Social media basado en GC a partir de elementos esenciales para un proceso de software, donde de forma detallada se describen las actividades asociadas a productos de trabajo de entrada y salida, y se sugieren un conjunto de herramientas para su aplicación en función a diversos roles.

5.2 Trabajo futuro

Como trabajo futuro, se destaca que en esta área de investigación existen diversas direcciones para seguir avanzando en el análisis de datos de redes sociales y la mejora del User Engagement, Interactividad y Viralización en las redes sociales. A continuación, se mencionan algunas:

- El ajustar los modelos de análisis estadístico y de regresión para mejorar la precisión de interpretación y la capacidad predictiva. Esto podría considerar la incorporación de técnicas de aprendizaje automático para adaptarse a los patrones de comportamiento de los usuarios.

- Cada plataforma cuenta con su propio conjunto de métricas y características. Por lo tanto, el realizar la integración del Marco de trabajo propuesto con otras plataformas de redes sociales lograría ampliar el alcance del estudio.
- Considerar incorporar a los modelos la segmentación de los usuarios en grupos demográficos o de interés. Esto permitirá un alcance más efectivo en los contenidos para diferentes usuarios.
- Evaluar las estrategias diseñadas para incrementar la participación de los usuarios con el contenido creando experimentos que controlen los formatos de publicación, la frecuencia con que se publica y los enfoques para obtener la atención de los usuarios.
- Aplicar el marco de trabajo en otras organizaciones e instituciones para validar hallazgos a través de colaboraciones externas y expertos.
- Desarrollar herramientas y aplicaciones basadas en los resultados de esta investigación. Esto para el beneficio de los profesionales y gestores de redes sociales.

REFERENCIAS

- Abiteboul, S., & Dowek, G. (2020). *The Age of Algorithms*. Cambridge University Press.
- Adinolfi, P., D'Avanzo, E., Lytras, M. D., Novo-Corti, I., & Picatoste, J. (2016). Sentiment Analysis to Evaluate Teaching Performance. *International Journal of Knowledge Society Research*, 7(4), 86–107. <https://doi.org/10.4018/ijksr.2016100108>
- Agusriandi, Sukaesih Sitanggang, I., & Hartono Wijaya, S. (2020). *Student Performance Based on Activity Log on Social Network and e-Learning*. 10(6).
- Al-Dheleai, Y. M., & Tasir, Z. (2017). Using facebook for the purpose of students' interaction and its correlation with students' academic performance. *Turkish Online Journal of Educational Technology*, 16(4), 170–178.
- Al-Maatouk, Q., Othman, M. S., Aldraiweesh, A., Alturki, U., Al-Rahmi, W. M., & Aljeraiwi, A. A. (2020). Task-technology fit and technology acceptance model application to structure and evaluate the adoption of Social media in academia. *IEEE Access*, 8, 78427–78440. <https://doi.org/10.1109/ACCESS.2020.2990420>
- Alles, M. (2012). *Social media y Recursos Humanos* (1st ed.). Ediciones Granica.
- Alsufyan, N. K., & Aloud, M. (2017). The state of Social media engagement in Saudi universities. *Journal of Applied Research in Higher Education*, 9(2), 267–303. <https://doi.org/10.1108/JARHE-01-2016-0001>
- Amin, A., & Rajadurai, J. (2018). The Conflict between Social media and Higher Education Institutions. *Global Business and Management Research: An International Journal*, 10(4).
- Arias, J. E., & Aristizábal, C. A. (2011). El dato, la información, el conocimiento y su productividad en empresas del sector público de Medellín. *Semestre Económico*, 14(28), 95–109. <https://doi.org/10.22395/seec.v14n28a6>
- Arora, A., Bansal, S., Kandpal, C., Aswani, R., & Dwivedi, Y. (2019). Measuring Social media influencer index- insights from facebook, Twitter and Instagram. *Journal of Retailing and Consumer Services*, 49(February), 86–101. <https://doi.org/10.1016/j.jretconser.2019.03.012>
- Asur, S., & Huberman, B. A. (2010). Predicting the future with Social media. *Proceedings - 2010 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2010*, 1, 492–499. <https://doi.org/10.1109/WI-IAT.2010.63>
- Ayhan, S., Pesce, J., Comitz, P., Sweet, D., Bliesner, S., & Gerberick, G. (2013). Predictive analytics with aviation big data. *The Integrated Communications, Navigation and Surveillance Conference (ICNS)*.
- Ballesteros Herencia, C. A. (2018). El índice de engagement en redes sociales, una medición emergente en la Comunicación académica y organizacional. *RAZÓN Y PALABRA Primera Revista Electrónica En Iberoamérica Especializada En Comunicación*, 22(3_102), 96–124.
- Batrinca, B., & Treleaven, P. C. (2015). Social media analytics: a survey of techniques, tools and platforms. *AI and Society*, 30(1), 89–116. <https://doi.org/10.1007/s00146-014-0549-4>
- Bedi, P., & Sharma, C. (2016). Community detection in social networks. In *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* (Vol. 6, Issue 3, pp. 115–135). Wiley-Blackwell. <https://doi.org/10.1002/widm.1178>
- Biswal, S., & Muley, P. (2018). Data Mining in Social Network Analysis. *International Journal of Innovations in Engineering Research and Technology*, 5(5).

- Bitiktas, F., & Tuna, O. (2020). Social media usage in container shipping companies: Analysis of Facebook messages. *Research in Transportation Business and Management*, 34(February). <https://doi.org/10.1016/j.rtbm.2020.100454>
- Bonsón, E., & Ratkai, M. (2013). A set of metrics to assess stakeholder engagement and social legitimacy on a corporate Facebook page. *Online Information Review*, 37(5), 787–803. <https://doi.org/10.1108/OIR-03-2012-0054>
- Brenes Carranza, J. A., Martínez Porras, A., Quesada López, C. U., & Jenkins Coronas, M. (2020). Sistemas de apoyo a la toma de decisiones que usan inteligencia artificial en la agricultura de precisión. *RISTI - Revista Ibérica de Sistemas y Tecnologías de La Información*, 28(18), 217–229. <http://risti.xyz/issues/ristie28.pdf>
- Buxarrais, M. R. (2016). Redes sociales y educación Social networks and education. *Education in the Knowledge Society*, 17(2), 15–20. <http://www.redalyc.org/pdf/2010/201046521002.pdf%0Ahttp://dx.doi.org/10.14201/eks20161721520>
- Campos-Freire, F., & Rúas-Araújo, R. (2016). Uso de las redes sociales digitales profesionales y científicas: el caso de las 3 universidades gallegas. *El Profesional de La Información*, 25(3), 431. <https://doi.org/10.3145/epi.2016.may.13>
- Celaya, J. (2008). *La empresa en la Web 2.0*. Ediciones Gestión 2000.
- Chan-Olmsted, S., Wolter, L. C., & Wang, R. (2017). Toward a multidimensional framework of media engagement: Conceptualizing consumer experience and connection with media content in a digital environment. *Emma Conf 2017. European Media Management Association.*, 585.
- Chen, X., Madhavan, K., & Vorvoreanu, M. (2013). A web-based tool for collaborative Social media data analysis. *Proceedings - 2013 IEEE 3rd International Conference on Cloud and Green Computing, CGC 2013 and 2013 IEEE 3rd International Conference on Social Computing and Its Applications, SCA 2013*, 383–388. <https://doi.org/10.1109/CGC.2013.67>
- Dabbagh, N., & Kitsantas, A. (2012). Personal Learning Environments, Social media, and self-regulated learning: A natural formula for connecting formal and informal learning. *Internet and Higher Education*, 15(1), 3–8. <https://doi.org/10.1016/j.iheduc.2011.06.002>
- Dey, P., & Roy, S. (2016). A comparative analysis of different social network parameters derived from facebook profiles. *Advances in Intelligent Systems and Computing*, 379, 125–132. https://doi.org/10.1007/978-81-322-2517-1_13
- Diliana, N. A., & Indrawati. (2022). Identification of Influencers and Community of Lazada Using Social Network Analysis. *2022 International Conference on Data Science and Its Applications, ICoDSA 2022*, 305–308. <https://doi.org/10.1109/ICoDSA55874.2022.9862849>
- El Haddaoui, B., Chiheb, R., Faizi, R., & Afia, A. El. (2018, May 2). Toward a sentiment analysis framework for Social media. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3230905.3230919>
- Elkaseh, A. M., Wong, K. W., & Fung, C. C. (2016). Perceived ease of use and perceived usefulness of Social media for E-Learning in libyan higher education: A structural equation modeling analysis. *Int. J. Inf. Edu. Technol.*, 6(3), 192–199.
- Erciyes, K. (2021). Algorithms. In: *Discrete Mathematics and Graph Theory. Undergraduate Topics in Computer Science*. Springer, Cham.
- Eriksson, N., Sjöberg, A., Rosenbröijer, C.-J., & Fagerstrøm, A. (2019). Consumer brand post engagement on Facebook and Instagram-A study of three interior design brands. *International Conference on*

Electronic Business (ICEB) at AIS Electronic Library (AISeL), 12–20.
<https://aisel.aisnet.org/iceb2019>

- Flores-Rios, B. L. (2016). Modelo de evolución de la gestión del conocimiento en MiPyMes, de acuerdo con el nivel de madurez en un programa de mejora de procesos de software. In *Universidad Autónoma de Baja California* (Issue October 2016).
https://www.researchgate.net/publication/332058888_Modelo_de_evolucion_de_la_gestion_del_conocimiento_en_MiPyMes_de_acuerdo_con_el_nivel_de_madurez_en_un_programa_de_mejora_de_procesos_de_software
- Froment, F., García-González, A. J., & Cabero, J. (2022). The relationship of Twitter with teacher credibility and motivation in university students. *Comunicar*, 30(71), 1–12.
<https://doi.org/10.3916/C71-2022-10>
- Froment, F., García González, A. J., & Bohórquez, M. R. (2017). The use of social networks as a communication tool between teachers and students: A literature review. *Turkish Online Journal of Educational Technology*, 16(4), 126–144.
- Furtado Avanza, M., & Moreira Pinheiro, W. (2018). Audience Measurement in Press Through Social Interactivity on Facebook in Brazil. *Mediaciones*, 14(20), 48–69.
<https://doi.org/10.26620/uniminuto.mediaciones.14.20.2018.48-69>
- Ghani, N. A., Hamid, S., Targio Hashem, I. A., & Ahmed, E. (2019). Social media big data analytics: A survey. *Computers in Human Behavior*, 101(July 2018), 417–428.
<https://doi.org/10.1016/j.chb.2018.08.039>
- Gil-quintana, J., de León, E. V., Osuna-acedo, S., & Marta-lazo, C. (2022). Nano-Influencers Edutubers: Perspective of Centennial Generation Families in Spain. *Media and Communication*, 10(1), 247–258.
<https://doi.org/10.17645/mac.v10i1.4760>
- Gil-Quintana, J., Vida de León, E., Osuna-Acedo, S., & Marta-Lazo, C. (2022). Nano-Influencers Edutubers: Perspective of Centennial Generation Families in Spain. *Media and Communication*, 10(1), 247–258. <https://doi.org/10.17645/mac.v10i1.4760>
- Gole, S., & Tidke, B. (2015). A survey of big data in Social media using data mining techniques. *ICACCS 2015 - Proceedings of the 2nd International Conference on Advanced Computing and Communication Systems*, 5–10. <https://doi.org/10.1109/ICACCS.2015.7324059>
- González, A. de J., Joaquín, C. Z., & Collazos, C. A. (2009). Karagabi kmmodel: Modelo de referencia para la introducción de iniciativas de gestión del conocimiento en organizaciones basadas en conocimiento. *Ingeniare*, 17(2), 223–235. <https://doi.org/10.4067/s0718-33052009000200011>
- Gonzalez, C., & Munoz, L. (2016). Online Social networks its impact on higher Education: Case Study of Technological University of Panama. *Campus Virtuales*, 5(1), 84–90.
- Goyal, S. J., & Goyal, R. (2020). Social media Analytics – Challenges during the Data Analysis. *Artificial & Computational Intelligence*, 1(2). https://acors.org/ijacoi/VOL1_ISSUE2_10.pdf
- Gupta, V., & Gupta, M. (2016). Social media Mining: A New Framework and Literature Review. *International Journal of Business Analytics*, 3(1), 58–68. <https://doi.org/10.4018/IJBAN.2016010104>
- Han, J., Park, D., Forbes, H., & Schaefer, D. (2020). A computational approach for using social networking platforms to support creative idea generation. *Procedia CIRP*, 91, 382–387.
<https://doi.org/10.1016/j.procir.2020.02.190>
- Hanneman, R. A., & Riddle, M. (2009). *Introduction to Social Network Methods: Table of Contents*. [http://www.faculty.ucr.edu/~hanneman/nettext/\[8/](http://www.faculty.ucr.edu/~hanneman/nettext/[8/)
- He, Y., Wang, C., & Jiang, C. (2017). Mining Coherent Topics With Pre-Learned Interest Knowledge in

- Twitter. *IEEE Access*, 5, 10515–10525. <https://doi.org/10.1109/ACCESS.2017.2696558>
- Hernández-García, F., Robaina-Castillo, J. I., Herrera, L. A. L., Peña, J. A. S., Ávila, M. H., Pérez Calleja, N. C., Díaz, E. D. C. G., & Peraza, B. M. A. (2020). A digital platform to meet information needs of students' scientific activity in medical sciences schools. *Revista Cubana de Información En Ciencias de La Salud*, 31(2), 1–26. <https://doi.org/10.36512/rcics.v31i2.1455>
- Herrera, H. (2012). Social Networks: A new diffusion tool. *Reflexiones*, 91(2), 121–128. <https://revistas.ucr.ac.cr/index.php/reflexiones/article/view/1513/1521>
- Hill, R. K. (2015). What an Algorithm Is. *Philosophy & Technology*, 29(1), 35–59.
- Hu, Y. H., & Chen, K. (2016). Predicting hotel review helpfulness: The impact of review visibility, and interaction between hotel stars and review ratings. *International Journal of Information Management*, 36(6), 929–944. <https://doi.org/10.1016/j.ijinfomgt.2016.06.003>
- Huertas, A., Setó-Pàmies, D., & Míguez-González, M. (2015). Comunicación de destinos turísticos a través de los medios sociales. *El Profesional de La Información*, 24(1), 15–21. <https://doi.org/10.3145/epi.2015.ene.02>
- Humberstone, J. (2019). Análisis de redes sociales: Identificación de comunidades virtuales en Twitter. *Realidad y Reflexión*, 50(50), 70–81. <https://doi.org/10.5377/ryr.v50i50.9095>
- Injadat, M., Salo, F., Nassif, A. B., Injadat, M., Salo, F., & Nassif, A. B. (2016). Data Mining Techniques in Social media: A Survey. *Neurocomputing*. <http://dx.doi.org/10.1016/j.neucom.2016.06.045>
- Jayasingh, S. (2019). Consumer brand engagement in social networking sites and its effect on brand loyalty. *Cogent Business and Management*, 6(1). <https://doi.org/10.1080/23311975.2019.1698793>
- Karademir Coşkun, T., Erdoğan, F., & Kokoç, M. (2020). Students' Profiles of Information Seeking Behaviors on Facebook: Relations to Personality, Time Management, Gender, and Facebook Use. *International Journal of Technology in Education and Science (IJTES)*, 4(2), 144–159. www.ijtes.net
- Kaye, L. K. (2021). Exploring the “socialness” of Social media. *Computers in Human Behavior Reports*, 3, 100083. <https://doi.org/10.1016/j.chbr.2021.100083>
- Kim, C., & Yang, S. U. (2017). Like, comment, and share on Facebook: How each behavior differs from the other. *Public Relations Review*, 43(2), 441–449. <https://doi.org/10.1016/j.pubrev.2017.02.006>
- Kurnia, P. F., & Suharjito. (2018). Business Intelligence Model to Analyze Social media Information. *Procedia Computer Science*, 135, 5–14. <https://doi.org/10.1016/j.procs.2018.08.144>
- Kuz, A., Falco, M., & Giandini, R. (2016). Análisis de redes sociales: Un caso práctico. *Computacion y Sistemas*, 20(1), 89–106. <https://doi.org/10.13053/CyS-20-1-2321>
- Liao, S. H., Chu, P. H. y Hsiao, P. Y., (2012). Data mining techniques and applications – A decade review from 2000 to 2011. *Expert Systems with Applications*, 39(12), pp. 11303–11311.
- Liu, H., Morstatter, F., Tang, J., & Zafarani, R. (2016). The good, the bad, and the ugly: uncovering novel research opportunities in Social media mining. *International Journal of Data Science and Analytics*, 1(3–4), 137–143. <https://doi.org/10.1007/s41060-016-0023-0>
- Liyanage, C. R., & Premarathne, S. C. (2021). Clustered Approach for Clone Detection in Social media. *International Journal on Advanced Science, Engineering and Information Technology*, 11(1), 99–104. <https://doi.org/10.18517/ijaseit.11.1.9272>
- Martín Blanco, V. (2015). Social media: Herramientas de gestión para el community manager. In *Fundación Confemetal* (1ra ed., Vol. 1).
- Mazza, B., & Palermo, A. (2018). Social media content for business and user engagement on Facebook.

Journal for Communication Studies, 11(1), 49–73.

- Meng, X., Shi, L., Yao, L., Zhang, Y., & Cui, L. (2020). Herd Behavior in Social media: The Role of Facebook Likes, Strength of Ties, and Expertise. *Information & Management*, 11, 124658. <https://doi.org/10.1016/j.colsurfa.2020.124658>
- Milstein, S., Lorica, B., Magoulas, R., Hochmuth, G., Chowdhury, A., & O'Reilly, T. (2008). Twitter and the micro-messaging revolution: Communication, connections, and immediacy. *O'Reilly Media, Incorporated*.
- Molina, O. E., Reynaldo, D., & Cancell, F. (2021). Engagement and desertion in MOOCs : Systematic review Engagement and desertion in MOOCs : Systematic review El engagement y la deserción en los MOOCs : Revisión sistemática. *Comunicar*, October, 111–124.
- Morales-i-Gras. (2021). *Comprensión de las redes sociales* (FUOC). Fundació Universitat Oberta de Catalunya.
- Mushtaq, H., Malik, B. H., Shah, S. A., Siddique, U. Bin, Shahzad, M., & Siddique, I. (2018). Implicit and explicit knowledge mining of Crowdsourced communities: Architectural and technology verdicts. *International Journal of Advanced Computer Science and Applications*, 9(1), 105–111. <https://doi.org/10.14569/IJACSA.2018.090114>
- Natthaphong Phuntusil, & Yachai Limpiyakorn. (2017). Predicting engaging content for increasing organic reach on facebook. *Lecture Notes in Electrical Engineering*, 424, 637–644. https://doi.org/10.1007/978-981-10-4154-9_73
- Niciporuc, T. (2014). Comparative analysis of the engagement rate on Facebook and Google Plus social networks. *Proceedings of International Academic Conferences*.
- Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. Oxford university press.
- Olea Méndez, J. E., & Moscosa Ruiz, P. (2018). *Plataforma digital para servicios educativos en la Delegación Cuauhtémoc*.
- Orioque, J. A. (2021). Student use of Facebook groups as a support for academic learning. *International Journal of Applied Science and Engineering*, 18(4(Special Issue)), 1–8. [https://doi.org/10.6703/IJASE.202106_18\(4\).001](https://doi.org/10.6703/IJASE.202106_18(4).001)
- Ovadia, S. (2014). ResearchGate and Academia.edu: Academic Social Networks. *Behavioral and Social Sciences Librarian*, 33(3), 165–169. <https://doi.org/10.1080/01639269.2014.934093>
- Oviedo-García, M. Á., Muñoz-Expósito, M., Castellanos-Verdugo, M., & Sancho-Mejías, M. (2014). Metric proposal for customer engagement in Facebook. *Journal of Research in Interactive Marketing*, 8(4), 327–344. <https://doi.org/10.1108/JRIM-05-2014-0028>
- Penttinen, V., & Frösén, J. (2022). Shared analytical capabilities in business networks. *Industrial Marketing Management*, 106, 308-322
- Peruta, A., & Shields, A. B. (2017). Social media in higher education: understanding how colleges and universities use Facebook. *Journal of Marketing for Higher Education*, 27(1), 131–143. <https://doi.org/10.1080/08841241.2016.1212451>
- Petersen, K., Vakkalanka, S., & Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, 64, 1–18. <https://doi.org/10.1016/j.infsof.2015.03.007>
- Pilař, L., Kvasničková Stanislavská, L., Kvasnička, R., Bouda, P., & Pítrová, J. (2021). Framework for Social media analysis based on hashtag research. *Applied Sciences (Switzerland)*, 11(8).

<https://doi.org/10.3390/app11083697>

- Prada Núñez, R., Hernández Suárez, C. A., & Aloiso Gamboa, A. (2019). Usos y efectos de la implementación de una plataforma digital en el proceso de enseñanza de futuros docentes en matemáticas. *Revista Virtual Universidad Católica Del Norte*, 5821(57), 137–156. <https://doi.org/10.35575/rvucn.n57a10>
- Prestridge, S., Jacobsen, M., Mulla, S., Paredes, S. G., & Charania, A. (2021). New alignments for the digital age: insights into connected learning. *Educational Technology Research and Development*, 69(4), 2171–2186. <https://doi.org/10.1007/s11423-021-09968-5>
- Purba, K. R., Asirvatham, D., & Murugesan, R. K. (2020). An analysis and prediction model of outsiders percentage as a new popularity metric on Instagram. *ICT Express*, 6(3), 243–248. <https://doi.org/10.1016/j.icte.2020.07.001>
- Purohit, H., Castillo, C., & Pandey, R. (2020). Ranking and grouping Social media requests for emergency services using serviceability model. *Social Network Analysis and Mining*, 10(1). <https://doi.org/10.1007/s13278-020-0633-3>
- Quiñonez Zuñiga, C., & Rivera Martinez, W. F. (2021). Modelo de gestión del conocimiento para centros de productividad e innovación. *TELOS: Revista de Estudios Interdisciplinarios En Ciencias Sociales*, 23(2), 347–366.
- Ramírez-Alpízar, A., Jenkins, M., Martínez, A., & Quesada-Lopéz, C. (2020). Uso de técnicas de minería de datos y aprendizaje automático para la detección de fraudes en estados financieros: un mapeo sistemático de literatura. *RISTI - Revista Ibérica de Sistemas y Tecnologías de La Información*, 28(April), 97–109. <https://www.researchgate.net/publication/340654299>
- Rodriguez-Vazquez, C., Teijeiro-Alvarez, M., & Blazquez-Lozano, F. (2016). La educación superior en España y las redes sociales: Instituciones públicas versus privadas. *Iberian Conference on Information Systems and Technologies, CISTI, 2016-July*, 69–92. <https://doi.org/10.1109/CISTI.2016.7521596>
- Rousidis, D., Koukaras, P., & Tjortjis, C. (2020). Social media prediction: a literature review. *Multimedia Tools and Applications*, 79(9–10), 6279–6311. <https://doi.org/10.1007/s11042-019-08291-9>
- Sabate, F., Berbegal-Mirabent, J., Cañabate, A., & Lebherz, P. R. (2014). Factors influencing popularity of branded content in Facebook fan pages. *European Management Journal*, 32(6), 1001–1011. <https://doi.org/10.1016/j.emj.2014.05.001>
- Salloum, S. A., Al-Emran, M., Monem, A. A., & Shaalan, K. (2017). A survey of text mining in Social media: Facebook and Twitter perspectives. *Advances in Science, Technology and Engineering Systems*, 2(1), 127–133. <https://doi.org/10.25046/aj020115>
- Sánchez-Rodríguez, J., Ruiz Palmero, J., & Sánchez Rivas, E. (2015). Problematic use of social networks in university students. *Revista Complutense de Educacion*, 26(2015), 159–174. https://doi.org/10.5209/rev_RCED.2015.v26.46360
- Savelev, A. O., Karpova, A. Y., Chaykovskiy, D. V., Vilnin, A. D., Kaida, A. Y., Kuznetsov, S. A., Igumnov, L. O., & Maksimova, N. G. (2021). The high-level overview of Social media content search engine. *IOP Conference Series: Materials Science and Engineering*, 1019(1). <https://doi.org/10.1088/1757-899X/1019/1/012097>
- Seaver, N. (2019). Knowing Algorithms. *DigitalSTS*, (February), 412–422. <https://doi.org/10.2307/j.ctvc77mp9.30>
- Shahbaznezhad, H., Dolan, R., & Rashidirad, M. (2021). The Role of Social media Content Format and Platform in Users' Engagement Behavior. *Journal of Interactive Marketing*, 53, 47–65. <https://doi.org/10.1016/j.intmar.2020.05.001>

- Shang, S. S. C., Li, E. Y., Wu, Y. L., & Hou, O. C. L. (2011). Understanding Web 2.0 service models: A knowledge-creating perspective. *Information and Management*, 48(4–5), 178–184. <https://doi.org/10.1016/j.im.2011.01.005>
- Sharma, K., & Lulandala, E. E. (2021). Facebook Ad Engagement: A Cross-cultural Analysis. *Global Business Review*. <https://doi.org/10.1177/09721509211007115>
- Simpao, A. F., Ahumada, L. M., Gálvez, J. A., & Rehman, M. A. (2014). A review of analytics and clinical informatics in health care. *Journal of Medical Systems*, 38(4), 45.
- Singh, N., Jaiswal, A., & Singh, T. (2022). Best time to post and review on Facebook and Instagram: analytical evidence. *South Asian Journal OfMarketing*. <https://doi.org/10.1108/SAJM-09-2022-0059>
- Singson, M., & Anees, M. (2017). Use of ResearchGate by the research scholars of Pondicherry University: A study. *DESIDOC Journal of Library and Information Technology*, 37(5), 366–371. <https://doi.org/10.14429/djlit.37.11139>
- Statista. (2022, June). *Redes sociales con mayor número de usuarios activos a nivel mundial en enero de 2022*. Social media y Contenido Creado Por El Usuario.
- Statista. (2023, Diciembre). *Redes sociales con mayor número de usuarios activos a nivel mundial en enero de 2023*. Social media y Contenido Creado Por El Usuario.
- Steinmetz, C., Rahmat, H., Marshall, N., Bishop, K., Thompson, S., Park, M., Corkery, L., & Tietz, C. (2020). Liking, Tweeting and Posting: An Analysis of Community Engagement through Social media Platforms. *Urban Policy and Research*, 1–21. <https://doi.org/10.1080/08111146.2020.1792283>
- Stieglitz, S., Mirbabaie, M., Ross, B., & Neuberger, C. (2018). Social media analytics – Challenges in topic discovery, data collection, and data preparation. *International Journal of Information Management*, 39(October 2017), 156–168. <https://doi.org/10.1016/j.ijinfomgt.2017.12.002>
- Surya Gunawan, T., Aleah Jehan Abdullah, N., Kartiwi, M., & Ihsanto, E. (2020). Social Network Analysis using Python Data Mining. *2020 8th International Conference on Cyber and IT Service Management, CITSM 2020*. <https://doi.org/10.1109/CITSM50537.2020.9268866>
- Sutanto, S. M. (2016). *The Use of Facebook Fanpage as Promotional Media for the Novel Book fans by using the Model of SMCR Berlo*.
- Toharia, M. (2010). Políticas de comunicación universitaria y divulgación científica. *La Cuestión Universitaria*, 6, 95–102.
- Tsinovoi, A. (2020). The management of visibility in digital diplomacy: Infrastructures and techniques. *First Monday*, May. <https://doi.org/10.5210/fm.v25i5.10116>
- Tsvetkov, V. Y. (2021). Qualitative analysis relations. *IOP Conference Series: Materials Science and Engineering*, 1047(1). <https://doi.org/10.1088/1757-899X/1047/1/012127>
- Ureña, R., Chiclana, F., Melançon, G., & Herrera-Viedma, E. (2019). A social network-based approach for consensus achievement in multiperson decision making. *Information Fusion*, 47(July 2018), 72–87. <https://doi.org/10.1016/j.inffus.2018.07.006>
- Velazquez-Solis. (2017). *Minería de procesos para la explotación de conocimiento en empresas de desarrollo de software con prácticas de CMMI-DEV*. Tesis de Maestría. Universidad Autónoma de Baja California, México.
- Velazquez-Solis, P., Flores-Rios, B. L., Astorga-Vargas, M. A., Ibarra-Esquer, J. E., González-Navarro, F. F., & Aguilar Vera, R. A. (2022). Analysis of scientific dissemination posts on Facebook from a Social media approach. *2022 17th Iberian Conference on Information Systems and Technologies (CISTI)* (pp. 1-6). IEEE.

- Velazquez-Solis, Paola E., Flores-Rios, B. L., Ibarra-Esquer, J. E., Astorga-Vargas, M. A., Aguilar Vera, R. A., & Ramírez-Barreto, Elizabeth López-Morteo, G. (2021). Identificación de áreas de aplicación de arquitecturas de software basadas en modelos, técnicas y herramientas de Social media. *Revista Ibérica de Sistemas e Tecnologías de Informação (RISTI)*, (42), 12-29.
- Velazquez-Solis, Paola Elvira, & Flores-Rios, B. L. (2021). Knowledge Flows Approach for Design of a Social media Software Architecture. *In 2021 Mexican International Conference on Computer Science (ENC)*, México. (pp. 1-8). IEEE.
- Verhoef, P. C., & Lemon, K. N. (2013). Successful customer value management: Key lessons and emerging trends. *European Management Journal*, 31(1), 1–15. <https://doi.org/10.1016/j.emj.2012.08.001>
- Voorveld, H. A. M., van Noort, G., Muntinga, D. G., & Bronner, F. (2018). Engagement with Social media and Social media Advertising: The Differentiating Role of Platform Type. *Journal of Advertising*, 47(1), 38–54. <https://doi.org/10.1080/00913367.2017.1405754>
- Wahyuningtyas, R., Alamsyah, A., & Diliana, N. A. (2021). Mapping Digital Talent Based on Competency using Social Network Analysis. *Proceeding - 2021 2nd International Conference on ICT for Rural Development, IC-ICTRuDev 2021*. <https://doi.org/10.1109/IC-ICTRuDev50538.2021.9656521>
- Wang, G., Gunasekaran, A., Ngai, E. W., & Papadopoulos, T. (2016). Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *International Journal of Production Economics*, 176, 98–110.
- Wheatley, D., & Vatnoey, E. (2020). ‘It’s Twitter, a bear pit, not a debating society’: A qualitative analysis of contrasting attitudes towards Social media blocklists. *New Media and Society*, 22(1), 5–25. <https://doi.org/10.1177/1461444819858278>
- Yaakop, A., Mohamed Anuar, M., & Omar, K. (2013). Like it or not: Issue of credibility in Facebook advertising. *Asian Social Science*, 9(3), 154–163. <https://doi.org/10.5539/ass.v9n3p154>
- Younis, E. (2015). Sentiment Analysis and Text Mining for Social media Microblogs using Open Source Tools: An Empirical Study. *International Journal of Computer Applications*, 112(5), 975–8887. <https://www.researchgate.net/publication/272463313>
- Zeng, D., Chen, H., Lusch, R., & Li, S. H. (2010). Social media analytics and intelligence. *IEEE Intelligent Systems*, 25(6), 13–16. <http://dx.doi.org/10.1109/MIS.2010.151>.

Anexo 1. Lista de productos académicos

A continuación, se detallan los productos académicos llevados a cabo en este proyecto. En cada uno de los elementos presentados, se enumeran las aportaciones correspondientes a cada capítulo.

En el capítulo II, destinado a presentar un contexto de los diferentes elementos relacionados, se llevó a cabo la investigación de los fundamentos teóricos y conceptuales esenciales sobre los cuales se cimienta la contribución de conocimiento al campo del Social media.

¿Cuántas mujeres se han graduado de un posgrado en química, física y medio ambiente a nivel nacional?

Capítulo de libro para Miscelánea Científica en México (2020). El objetivo era contextualizar el sujeto de estudio usando como fuente de información los anuarios de la Asociación Nacional de Universidades e Instituciones en Educación Superior (ANUIES), para realizar un análisis descriptivo, con perspectiva de género, considerando el nivel de participación de las mujeres que cuentan con estudios de posgrado en Instituciones de Educación Superior mexicanas, específicamente en las áreas disciplinarias de Ingeniería y Tecnología, Ciencias Naturales y Exactas y Ciencias de la Salud.

Identificación de áreas de aplicación de arquitecturas de software basadas en modelos, técnicas y herramientas de Social media.

Artículo indexado (Q4) en la Revista Ibérica de Sistemas e Tecnologías de Informação (2021). El objetivo fue identificar los enfoques, modelos de análisis, técnicas y herramientas de Social media aplicados en diferentes enfoques usando redes sociales, con el fin de detectar las áreas de aplicación por medio de estudios empíricos.

En el capítulo III, buscando llegar a la propuesta de marco de trabajo para el minado de Social media basado en Gestión de Conocimiento, se aplicó una metodología de intervención para tener mayor conocimiento sobre procesos de Social media y roles involucrados.

Dimensión epistemológica aplicada a procesos de Social media de un grupo académico de Facebook.

Artículo indexado en la revista electrónica de Computación, Informática, Biomédica y Electrónica (2021). Se empleó la metodología KoFI para identificar los elementos involucrados en la dimensión epistemológica y establecer la trazabilidad entre los flujos de conocimiento con los procesos de Social media y las herramientas tecnológicas que se utilizan para descubrir, recopilar, preparar y analizar los datos.

En el capítulo IV, se presentan los resultados de tres experimentos con muestras de datos diferentes, aplicando la propuesta de marco de trabajo presentada en el capítulo III.

Uso de métricas de interactividad para la interpretación de publicaciones de una fan page institucional en Facebook.

Artículo indexado en la revista electrónica de Computación, Informática, Biomédica y Electrónica (2023). Se realizó un análisis cuantitativo y cualitativo del Engagement, visto a través de la perspectiva de la universidad con su comunidad. Los resultados de la investigación ofrecen una herramienta estratégica para mejorar la difusión de contenido científico orientado a la participación de los usuarios.

Identificación de una comunidad universitaria en Facebook para la Divulgación de la Ciencia y la Cultura

Artículo indexado en la Revista Ingenierías Universidad de Medellín (2023). El objetivo principal era caracterizar la comunidad virtual de una Institución de Educación Superior. Para ello, se usaron los algoritmos Fruchterman-Reingold, Yifan Hu y Noverlap, de la herramienta Gephi, para el análisis y modelado de 30 fan pages de Facebook dedicadas a la promoción de la cultura y divulgación de la ciencia.

A regression model for analysis of User Engagement on Facebook fan page for scientific dissemination content

Artículo JCR (Q4) en Programming and Computer Software (2023). Se utilizaron métodos de análisis estadístico y análisis cualitativo para establecer un nuevo método para calcular el User Engagement y la interactividad en fan pages de Facebook enfocado en la difusión de información

científica. Usando los procesos de Social media, evaluación de métricas por coeficientes de correlación de Spearman y categorización de publicaciones por tipo de formato y fuente de contenido.

Anexo 2. Participación en eventos académicos

Se realizaron diferentes participaciones en eventos académicos, los cuales se enlistan a continuación:

- XVII Encuentro de la participación de la mujer en la Ciencia (2020). León, Guanajuato, México.
- Encuentro Nacional de Computación (2021). Morelia, Michoacán, México.
- 17ª Conferencia Ibérica de Sistemas y Tecnologías de Información (2022). Madrid, España
- Encuentro Nacional de Computación (2023). Ciudad de Guanajuato, Guanajuato, México.
- Jornadas Iberoamericanas de Ingeniería de Software e Ingeniería del Conocimiento (2023)