



UNIVERSIDAD LAICA “ELOY ALFARO” DE MANABÍ

EXTENSIÓN EL CARMEN

CARRERA EN INGENIERÍA EN TECNOLOGÍAS DE LA INFORMACIÓN.

Creada Ley No. 10 – Registro Oficial 313 de noviembre 13 de 1985

**PROYECTO INTEGRADOR PREVIO A LA OBTENCIÓN DEL TÍTULO DE
INGENIERO EN TECNOLOGÍAS DE LA INFORMACIÓN**

TEMA:

Sistema Informático Con Minería De Datos Para La Difusión De Contenido En Las Redes
Sociales De La Uleam El Carmen

AUTOR:

CAGUA YOONG CRISTOPHER JOSEPH

TUTOR:

ARCA ZAVALA JEFFERSON OMAR

FECHA:

16/08/2026

PAGINA DE RESPETO

CERTIFICACIÓN DEL DIRECTOR DE TRABAJO DE GRADUACIÓN

	NOMBRE DEL DOCUMENTO:	CÓDIGO: PAT-04-F-004
	CERTIFICADO DE TUTOR(A).	
	PROCEDIMIENTO: TITULACIÓN DE ESTUDIANTES DE GRADO BAJO LA UNIDAD DE INTEGRACIÓN CURRICULAR	REVISIÓN: 1 Página 1 de 1

CERTIFICACIÓN

En calidad de docente tutor de la Extensión El Carmen de la Universidad Laica "Eloy Alfaro" de Manabí, CERTIFICO:

Haber dirigido, revisado y aprobado preliminarmente el Trabajo de Integración Curricular bajo la autoría del estudiante **CAGUA YOONG CRISTOPHER JOSEPH**, legalmente matriculado en la carrera de Ingeniería en Software, período académico 2026-1, cumpliendo el total de 384 horas, cuyo tema del proyecto es **"SISTEMA INFORMÁTICO CON MINERÍA DE DATOS PARA LA DIFUSIÓN DE CONTENIDO EN LAS REDES SOCIALES DE LA ULEAM EL CARMEN"**.

La presente investigación ha sido desarrollada en apego al cumplimiento de los requisitos académicos exigidos por el Reglamento de Régimen Académico y en concordancia con los lineamientos internos de la opción de titulación en mención, reuniendo y cumpliendo con los méritos académicos, científicos y formales, y la originalidad del mismo, requisitos suficientes para ser sometida a la evaluación del tribunal de titulación que designe la autoridad competente.

Particular que certifico para los fines consiguientes, salvo disposición de Ley en contrario.

El Carmen, 04 de agosto de 2026

Lo certifico,


Ing. Jefferson Omar Arca Zavala, Mgs.
Docente Tutor
Área: Ingeniería en Software

TRIBUNAL DE SUSTENTACIÓN



Uleam

Extensión El Carmen

TRIBUNAL DE SUSTENTACIÓN

Título del trabajo de titulación:

Sistema Informático con Minería de Datos para la Difusión de Contenido en las Redes Sociales de la Uleam El Carmen

Modalidad:

Proyecto Integrador

Autor:

Cagua Yoong Cristopher Joseph

Tutor:

Ing. Arca Zavala Jefferson Omar, Mgtr

Tribunal de Sustentación

- Presidente A. S. Minaya Macias Renelmo Wladimir, Mgtr.

- Miembro Ing. Pozo Hernández Clara Guadalupe, Mgtr.

- Miembro Ing. Angel Wilson Villareal Cobeña, Mgtr

Fecha de sustentación:

20 agosto 2026

DECLARACIÓN DE AUTORÍA



La responsabilidad de todo el contenido de este trabajo de titulación, con el tema:

Sistema informático con minería de datos para la difusión de contenido en las redes sociales de la Uleam El Carmen, corresponde exclusivamente a: **Cagua Yoong Cristopher Joseph** con C.I. **1315447985** y los derechos patrimoniales de la misma corresponden a la Universidad Laica "Eloy Alfaro" de Manabí.

A handwritten signature in blue ink, appearing to read "Cagua Yoong", is written over a horizontal line.

Cagua Yoong Cristopher Joseph

C.I. 1315447985

DEDICATORIA

Esta dedicatoria se las otorgo a mis padres, con todo mi amor y gratitud. Gracias por cada sacrificio, por creer en mí incluso cuando yo mismo dudaba, por impulsarme a seguir adelante y por enseñarme que el esfuerzo y la perseverancia siempre tienen recompensa. Este logro también les pertenece, porque sin su apoyo incondicional no habría sido posible llegar hasta aquí.

A mis hermanos, quienes con sus ocurrencias, bromas, consejos y compañía hicieron que este camino fuera mucho más llevadero. Compartimos momentos de alegría, dificultades y aprendizajes que siempre llevaré en mi corazón.

A mis tías y a todas las personas de mi familia que, de una u otra manera, me brindaron su apoyo, confianza y una mano cuando más lo necesité. Su ayuda fue fundamental para que nunca perdiera de vista este sueño.

A mis amigos, aquellos que estuvieron presentes en distintas etapas de mi vida, compartiendo experiencias, palabras de aliento y momentos inolvidables. Aunque hoy algunos caminos sean diferentes, siempre recordaré el apoyo que me brindaron durante este proceso.

Y, de manera muy especial, a mi novia. Gracias por permanecer a mi lado en los buenos y malos momentos, por creer en mí incluso cuando yo no veía mis propias capacidades, por tu paciencia, comprensión y amor incondicional. Gracias por acompañarme en este camino y por recordarme, cada día, que nunca se deja de crecer cuando se tiene a la persona correcta al lado.

A todos ustedes, gracias. Este logro no lleva únicamente mi nombre, también lleva el esfuerzo, el cariño y el apoyo de cada persona que formó parte de este camino. Con profunda gratitud, les dedico este trabajo y este nuevo comienzo profesional.

AGRADECIMIENTO

Expreso mi más sincero agradecimiento a Dios, por brindarme la fortaleza, la sabiduría y la perseverancia necesarias para culminar esta importante etapa de mi vida profesional.

A la Universidad Laica "Eloy Alfaro" de Manabí, Extensión El Carmen, por abrirme sus puertas y ofrecerme una formación académica de calidad, permitiéndome desarrollar los conocimientos y competencias que hoy hacen posible este logro.

A todos los docentes de la carrera de Ingeniería en Tecnologías de la Información, quienes, con su experiencia, dedicación y compromiso, compartieron sus conocimientos y contribuyeron significativamente a mi formación profesional y personal. Cada enseñanza impartida durante estos años será una herramienta valiosa para mi futuro.

Mi más profundo agradecimiento a mi familia, especialmente a mis padres y hermanos, por su amor, apoyo incondicional, paciencia y confianza durante todo este proceso. Ellos fueron mi mayor motivación para seguir adelante y nunca rendirme ante las dificultades.

A mis tías, familiares y amigos, quienes estuvieron presentes con una palabra de aliento, un consejo o un gesto de apoyo en los momentos en que más lo necesité. Su compañía y confianza fueron fundamentales para alcanzar esta meta.

Finalmente, agradezco a todas las personas que, de una u otra manera, contribuyeron al desarrollo de este trabajo de titulación y formaron parte de este importante capítulo de mi vida. A cada uno de ustedes, mi más sincera gratitud por acompañarme en este camino y ser parte de este logro, que representa el resultado del esfuerzo, la dedicación y el apoyo recibido a lo largo de mi formación universitaria.

ÍNDICE DE CONTENIDOS

Certificación del director de trabajo de graduación	¡Error! Marcador no definido.
Tribunal de sustentación	3
Declaración expresa de autoría	5
Dedicatoria	6
Agradecimiento	7
Índice de contenidos	8
Capítulo I	17
1 Introducción	17
1.1 Introducción	17
1.2 Presentación del tema	17
1.3 Ubicación y contextualización de la problemática	18
1.3.1 Planteamiento del problema.....	18
1.3.2 Problematización.....	18
1.3.3 Génesis del problema.....	19
1.3.4 Estado actual del problema	19
1.4 Diagrama causa – efecto del problema	20
1.5 Objetivos	20
1.5.1 Objetivo general.....	20
1.5.2 Objetivos Específicos.....	20
1.6 Justificación	21
1.7 Impactos esperados	21
1.7.1 Impactos tecnológicos.....	21
1.7.2 Impacto social	22

1.7.3	Impacto ecológico	22
Capítulo II		23
2	Marco teórico de la investigación	23
2.1	Antecedentes Históricos.....	23
2.1.1	Minería de datos.....	23
2.1.2	Sistemas Informáticos	23
2.1.3	Evolución de la comunicación y la necesidad de análisis en el entorno digital 24	
2.1.4	Origen y Evolución de la Minería de Datos.....	24
2.1.5	Las Redes Sociales como Canales Institucionales: De la interacción social al compromiso organizacional	24
2.2	Antecedente de investigación relacionados al tema presentado	24
2.2.1	Métricas de interacción y compromiso en redes sociales de universidades ecuatorianas 24	
2.2.2	Algoritmos de predicción de engagement.....	25
2.2.3	Minería de datos aplicada a la comunicación digital: Casos de éxito en otras instituciones educativas	25
2.2.4	Comportamiento del consumidor digital joven: Investigaciones sobre los hábitos de la Gen Z en Ecuador	25
2.3	Definiciones conceptuales	26
2.3.2	Minería de Datos y el Proceso KDD.....	27
2.3.3	Inteligencia Artificial y Machine Learning.....	27
2.3.4	Sistemas Informáticos: El núcleo de la solución	28
2.3.5	Análisis de Sentimientos.....	28
2.3.6	Visualización de Datos (Data Visualization).....	29
2.3.7	Inteligencia de Negocios.....	30

2.3.8	Metodologías de Desarrollo de Software y el Modelo en Cascada (Waterfall)	
	30	
2.4	Conclusión	31
Capítulo III.....		32
3	MARCO INVESTIGATIVO	32
3.1	Introducción	32
3.2	Tipo de investigación.....	32
3.2.1	Investigación Aplicada.....	32
3.2.2	Investigación Descriptiva.....	33
3.2.3	Investigación Correlacional	33
3.3	Métodos de investigación	34
3.3.1	Método Cuantitativo	34
3.3.2	Método Descriptivo	34
3.3.3	Método Analítico	34
3.4	Fuentes de información.....	35
3.4.1	Encuestas Online.....	35
3.5	Estrategia operacional para la recolección de datos	36
3.5.1	Población.....	36
3.5.2	Análisis de herramientas de recolección.....	37
3.5.3	Plan de recolección de datos	37
3.6	Análisis de resultados	38
3.6.1	Tabulación y análisis de los datos.....	38
3.6.2	Presentación y descripción de los resultados obtenido	40
3.6.3	Informe final del análisis de los datos.....	42
Capítulo IV.....		43

4	Marco propositivo (Elaboración de la propuesta).....	43
4.1	Introducción	43
4.2	Descripción de la propuesta	44
4.3	Determinación de recursos.....	44
4.3.1	Humanos	44
4.3.2	Tecnológicos	45
	Económico	47
4.4	Desarrollo.....	47
4.4.2	Diseño	53
4.4.3	Fase III	63
4.4.4	Pruebas del Sistema y Verificación de Software	72
	Capítulo V.....	76
5	EVALUACIÓN DE RESULTADOS	76
5.1	Introducción	76
5.2	Presentación y Monitoreo de Resultados	77
5.3	Interpretación Objetiva	78
	Capítulo VI.....	81
6	CONCLUSIONES.....	81
6.1	Conclusiones	81
7	Recomendaciones	82
	BIBLIOGRAFIA	84
	ANEXOS	89
	GLOSARIO	91

INDICE DE TABLAS

Tabla 1: Recolección de datos	37
Tabla 2: Tabulación y análisis de datos de encuesta.....	38
Tabla 3	40
Tabla 4: Roles y funciones de los actores estratégicos del sistema	45
Tabla 5: Determinación de recurso humano	45
Tabla 6: Recursos de hardware	46
Tabla 7: Recursos de software y especificaciones.	46
Tabla 8: Recursos económicos.....	47
Tabla 9: Tipos y roles de usuarios	52
Tabla 10: Resumen de la arquitectura de componentes y sus responsabilidades	66
Tabla 11: Casos de Prueba de Caja Blanca para la Lógica Interna del Backend.....	73
Tabla 12: Casos de Prueba de Caja Negra para la Interfaz de Escritorio en Electron.....	74
Tabla 13: Casos de Prueba de Integración del Pipeline de Datos Multi-Plataforma	75
Tabla 14: Monitoreo Comparativo de Indicadores Antes y Después de la Implementación de SoftMedia.....	77

INDICE DE ILUSTRACIONES

Ilustración 1: Diagrama causa-efecto del problema.....	20
Ilustración 4: Funciones del administrador en el Dashboard.....	53
Ilustración 5: Funciones del motor de Inteligencia Artificial	54
Ilustración 6: Arquitectura del motor de datos	55
Ilustración 7: Visualización del Dashboard interactivo y Minería automatizada y clasificación	56
Ilustración 8: Transición de los datos extraídos y Ciclo de ejecución de los mineros.....	57
Ilustración 9: Diagrama de base de datos: Entidad-Relación	58
Ilustración 10: Dashboard Institucional y Nivel de Aceptación Global	59
Ilustración 11: Módulo de Minería y Gestión de Publicaciones.....	60
Ilustración 12: Módulo de Análisis de Sentimientos (NLP).....	61
Ilustración 13: Estructura de Navegación Lateral (Sidebar).....	62
Ilustración 14: Módulo de Configuración de Persistencia y Conexión Relacional.	67
Ilustración 15: Modelos de Datos de Entidades Relacionales	68
Ilustración 16: Script Orquestador para la Extracción Automatizada de Redes Sociales. 69	
Ilustración 17: Consulta de métricas de interacción	70
Ilustración 18: Generación de token	71
Ilustración 19: Petición de publicaciones, reacciones y comentarios desde la Graph API de Facebook	71
Ilustración 20: Consulta de contenido multimedia e interacciones para Instagram Profesional	72

ÍNDICE DE ANEXO

ANEXOS	89
Anexo 1: Reporte del sistema antiplagio Compilatio: informe de similitud del trabajo donde muestra posibles coincidencias, su propósito, garantizar la originalidad del proyecto	89
Anexo 2: Evidencia de encuesta	90
Anexo 3: Evidencia de respuestas de estudiantes y personal administrativos.....	91

RESUMEN

Este estudio analiza la falta de eficacia en la difusión de información oficial dentro de la Universidad Laica Eloy Alfaro de Manabí, Extensión El Carmen. Con el propósito de mitigar esta falencia, la investigación plantea el diseño y despliegue de una plataforma informática basada en técnicas de minería de datos para optimizar el alcance de las publicaciones gestionadas por el equipo SoftMedia. Metodológicamente, el estudio se desarrolló bajo un enfoque cuantitativo de alcance descriptivo y diseño no experimental transversal; como técnica de recolección de datos, se aplicó una encuesta mediante un cuestionario estructurado a una muestra representativa de 2.074 personas (correspondiente al 33,3 % de la comunidad universitaria). Los hallazgos demostraron un perfil mayoritariamente estudiantil (96 %), concentrado en un rango de edad entre 18 y 24 años (81 %). En este grupo, el 71 % prefiere contenidos multimodales con audio y subtítulos, y el 39 % prioriza formatos audiovisuales inferiores a 60 segundos. Frente a este panorama, la arquitectura implementa módulos de inteligencia artificial para optimizar la toma de decisiones y programar publicaciones en horarios de mayor tráfico en Facebook, sustituyendo la gestión empírica por un modelo fundamentado en analítica de datos.

ABSTRACT

This study examines the ineffectiveness of disseminating official information within the Laica Eloy Alfaro de Manabí University, El Carmen Extension. To address this problem, the research proposes developing an information platform using data mining techniques to improve the reach of publications managed by the Softmedia team. To better understand user interaction within this digital ecosystem, a representative sample of 2,074 individuals covering 33.3% of the campus was analyzed. The findings showed a strong majority of students, who made up 96% of the total users, with 81% of the audience falling into the 18 to 24 age bracket. Within this group, 71% prefer multimedia content that includes audio and subtitles, while 39% limit their level of attention to a maximum of sixty seconds. This behaviour makes it clear that audience trends determine message retention, which means that institutional communication must be adapted to the real habits of the community. In light of this, the developed architecture uses artificial intelligence agents to optimise decisions, publishing vertical videos during the busiest night-time hours on Facebook. With this approach, the university moves away from empirical management and promotes a much more precise transfer of information, directly supported by data analytics. As a result, the adoption of this technological environment bridges the communication gap and ensures that organic dissemination flows in perfect harmony with the academic community.

CAPÍTULO I

2 INTRODUCCIÓN

2.1 Introducción

La Universidad Laica Eloy Alfaro de Manabí ha logrado construir una presencia muy fuerte en internet desde su sede principal en Manta. Con los avances en su infraestructura, la universidad estableció las redes sociales como canales formales de comunicación con la comunidad. Sin embargo, aún se requiere una plataforma centralizada para aplicar minería de datos y analizar los metadatos de las distintas sedes. Además, la ausencia de un sistema inteligente que mida las interacciones diarias dificulta el aprovechamiento de los datos reales, tanto en la sede principal como en extensiones como El Carmen.

Esta desconexión analítica evidencia que administrar canales de comunicación sin métricas precisas limita el alcance estratégico de la institución. Como resultado, la universidad desaprovecha su verdadero potencial informativo al no integrar todas sus sedes académicas bajo un mismo esquema de procesamiento digital.

En la extensión de El Carmen el escenario resulta todavía más complejo. Allí, la difusión de contenidos institucionales se topa con marcadas barreras tecnológicas y limitaciones de orden organizacional. Al contar con las herramientas de analítica web e indicadores de interacción, evaluar el impacto real de las publicaciones se vuelve una tarea sumamente difícil. Esta falta de trazabilidad provoca que mucha información relevante no llegue de manera oportuna ni clara a docentes y estudiantes. Es un vacío que castiga directamente la efectividad del mensaje. Como consecuencia de esta ceguera de datos, la presencia digital de la sede pierde fuerza; un factor que termina levantando barreras para construir un vínculo constante, fluido y cercano con toda la comunidad.

2.2 Presentación del tema

Sistema informático con minería de datos para la difusión de contenido en las redes sociales de la ULEAM El Carmen

2.3 Ubicación y contextualización de la problemática

La Universidad Laica Eloy Alfaro de Manabí (ULEAM), Extensión El Carmen, se localiza en el cantón El Carmen, al noroccidente de la provincia de Manabí, Ecuador. Esta unidad académica de educación superior pública constituye el principal polo formativo y de transferencia técnica en la zona norte manabita, articulando la formación universitaria con los sectores productivos y sociales de su área de influencia. La institución alberga una comunidad universitaria conformada por una población aproximada de 6.228 integrantes, integrada por el cuerpo docente, personal administrativo y un estamento estudiantil que representa el 96 % de la comunidad institucional. Su oferta académica de tercer nivel abarca carreras estratégicas orientadas a las demandas del entorno regional, entre las que destacan Tecnologías de la Información, Ingeniería Agropecuaria, Educación Básica, y Contabilidad y Auditoría.

Dentro de esta estructura, la divulgación de las actividades académicas, científicas, culturales e institucionales está a cargo del equipo de producción y gestión digital SoftMedia, el cual emplea redes sociales como Facebook, Instagram, TikTok y YouTube como sus principales canales de difusión. Sin embargo, la unidad carece de una infraestructura tecnológica que recopile, almacene y procese de manera automatizada las métricas de interacción generadas por los usuarios. La ausencia de herramientas de minería de datos impide determinar con precisión los patrones de consumo digital, las franjas horarias de mayor tráfico, los formatos de mayor retención y la percepción ciudadana manifestada en los comentarios. En consecuencia, la gestión comunicacional opera bajo un enfoque estrictamente empírico y manual, lo que reduce la efectividad en la entrega de información oficial, genera vacíos comunicativos y desaprovecha el impacto estratégico que la analítica de datos aportaría a la toma de decisiones institucionales.

2.3.1 Planteamiento del problema

2.3.2 Problematicación

Las redes sociales constituyen una herramienta fundamental e indispensable en el entorno universitario. Su utilidad trasciende la mera moda, consolidándose como el medio idóneo para la difusión rápida, directa y con amplio alcance de información académica, cultural y administrativa. Sin embargo, en la extensión El Carmen de la Universidad Laica

Eloy Alfaro de Manabí (ULEAM), a pesar de la presencia activa en estas plataformas, la efectividad y el impacto comunicacional esperado se ven limitados. De este modo, la institución pierde margen de maniobra; su capacidad para tomar decisiones respaldadas por estrategias sólidas queda drásticamente reducida cuando intenta difundir contenido en el entorno digital.

Carecer de mecanismos para medir la reacción de las personas ante las publicaciones genera bastante incertidumbre en la gestión diaria. Este vacío metodológico obstaculiza la identificación precisa de los ejes temáticos con mayor nivel de aceptación dentro de la comunidad estudiantil. El impacto sobre la universidad es directo, ya que enfrenta serios obstáculos para consolidar su ecosistema digital, fomentar la participación de la comunidad y garantizar que la información fluya de manera oportuna. Sumado a esto, esta carencia analítica entorpece la toma de decisiones. Sin un panorama claro, resulta casi imposible diseñar estrategias o aplicar mejoras que respondan genuinamente a los intereses y necesidades de los usuarios.

2.3.3 Génesis del problema

El origen de esta problemática radica en la rápida evolución de las redes sociales, las cuales han pasado de ser un recurso secundario a consolidarse como un canal de comunicación esencial para las instituciones educativas. Estas plataformas se han establecido como el espacio principal de interacción, conversación y obtención de información para la comunidad universitaria, especialmente para los estudiantes.

Sin embargo, en la extensión El Carmen de la Universidad Laica Eloy Alfaro de Manabí, la difusión de información sigue haciéndose de manera bastante convencional. Tradicional, incluso. No se aprovecha un sistema informático que integre minería de datos, una herramienta que permitiría y esto es clave analizar las respuestas reales de los usuarios y optimizar el impacto de cada publicación.

2.3.4 Estado actual del problema

Actualmente en la sede de El Carmen, la publicación de contenido en redes sociales se realiza de forma manual. Este escenario operativo deja al descubierto la falta de una infraestructura formal enfocada en monitorear los canales institucionales. En la práctica, carecer

de un sistema sustentado en la minería de datos convierte en una tarea imposible la correcta interpretación de las métricas de alcance y la interacción real de los usuarios. A la larga, operar sin el aval de indicadores precisos debilita el impacto de los mensajes institucionales y termina desgastando el posicionamiento general de la extensión académica.

2.4 Diagrama causa – efecto del problema

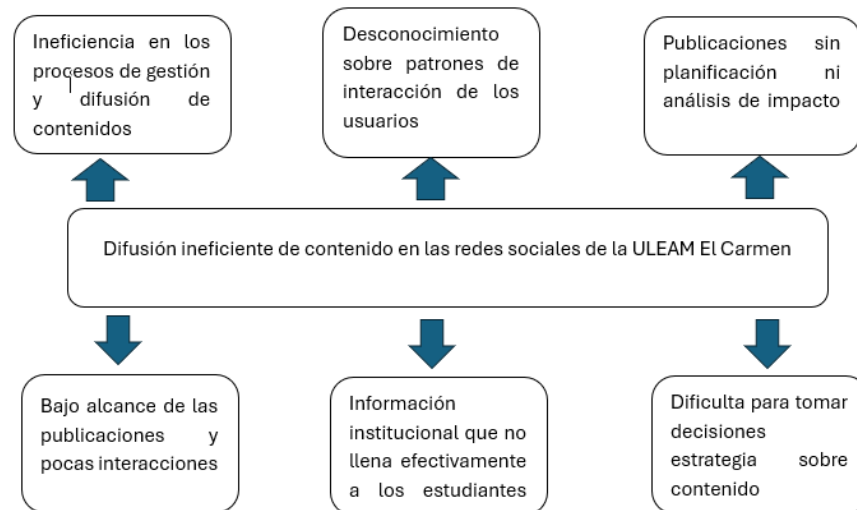


Ilustración 1: Diagrama causa-efecto del problema.

2.5 Objetivos

2.5.1 Objetivo general

Diseñar un sistema informático con minería de datos para la difusión de contenidos en redes sociales de la Uleam El Carmen

2.5.2 Objetivos Específicos

- ✓ Diagnosticar los factores técnicos y operativos que limitan la efectividad en la difusión de contenidos institucionales dentro de la planta central de la ULEAM El Carmen
- ✓ Sustentar bibliográficamente los fundamentos teóricos de la minería de datos y la comunicación digital para establecer el marco conceptual de la investigación.

- ✓ Examinar y comprender los patrones de consumo digital del colectivo universitario mediante el tratamiento y tabulación de las herramientas de recolección de datos administradas.
- ✓ Desarrollar la estructura de una solución informática que incorpore procedimientos de minería de datos para la mejora en la difusión de información institucional en plataformas digitales.
- ✓ Comprobar la eficacia de la solución planteada por medio de evaluaciones de operatividad que posibiliten cuantificar el aumento en la cobertura y la participación de los contenidos.

2.6 Justificación

Esta investigación surge de un apuro bien tangible: potenciar la forma en que se difunden contenidos en las redes sociales de la Universidad Laica Eloy Alfaro de Manabí, en su extensión de El Carmen. En estos tiempos, esas plataformas se han convertido en el pulso vital de la comunicación institucional, el lazo directo que conecta a la universidad con sus alumnos y el entorno que la rodea. Pero hay un tropiezo evidente. La ineficacia para hacer llegar la información digital ha tenido un impacto desfavorable: disminuye la presencia en el entorno digital, debilitan en el mapa general y, sobre todo, frena la involucración real de los estudiantes que integra esta comunidad educativa.

2.7 Impactos esperados

Este escenario operativo deja al descubierto la falta de una infraestructura formal enfocada en monitorear los canales institucionales. En la práctica, al no disponer de un sistema apoyado en la minería de datos, interpretar adecuadamente las métricas de alcance y la interacción real de los usuarios se vuelve una tarea imposible. Sin el respaldo de estas herramientas tecnológicas, el equipo encargado termina publicando a ciegas. En definitiva, se garantiza una comunicación digital más directa y segmentada, plenamente alineada con los estándares actuales de responsabilidad social y ambiental.

2.7.1 Impactos tecnológicos

Esta propuesta representa un paso clave para la ULEAM, ya que integra minería de datos para dejar atrás las suposiciones en la gestión de sus comunicaciones. Al migrar hacia un modelo sustentado en conocimiento verificado, la plataforma utiliza inteligencia

artificial aplicada para interpretar con precisión las expectativas de la comunidad. De esta manera, resulta mucho más sencillo entender al público y fomentar su participación activa en el entorno universitario. Asimismo, esta herramienta analítica respalda la toma de decisiones mediante evidencias objetivas, lo que otorga mayor seguridad a cada paso que da la institución. En consecuencia, la implementación de este entorno tecnológico dinamiza la interacción digital y afianza la proyección estratégica de la universidad.

2.7.2 Impacto social

Implementar plataformas que transmitan datos en tiempo real refuerza la transparencia institucional y la gobernanza universitaria. Mediante un control técnico que atenúa la dispersión informativa, la infraestructura asegura la entrega exacta de los contenidos clave a sus destinatarios. Este procedimiento ayuda a agilizar de forma notable los trámites administrativos del día a día. De esta manera, se garantiza que la información conserve su integridad y llegue siempre en los plazos previstos.

2.7.3 Impacto ecológico

Modernizar la infraestructura tecnológica e incorporar herramientas informáticas impulsa la transición hacia entornos verdaderamente digitales. Al favorecer la transmisión síncrona de contenidos, se suprime la necesidad de fabricación física masiva, posibilitando un manejo técnico eficaz que equipara la operatividad tecnológica con los criterios de sostenibilidad y responsabilidad ambiental.

CAPÍTULO II

3 MARCO TEÓRICO DE LA INVESTIGACIÓN

3.1 Antecedentes Históricos

3.1.1 Minería de datos

La evolución histórica de la minería de datos se articula a través de una trayectoria cuatrfásica que trasciende desde los fundamentos estadísticos clásicos hasta la complejidad computacional contemporánea. Según Chooramun (2025), los cimientos de esta disciplina se establecieron mediante los aportes pioneros de Bayes y Gauss en los siglos XVIII y XIX, proporcionando la base teórica necesaria para la posterior integración de redes neuronales artificiales y el procesamiento electrónico de datos a mediados del siglo XX. Un hito estructural en esta progresión ocurrió en la década de 1990 con la formalización del proceso de Descubrimiento de Conocimiento en Bases de Datos (KDD) y la creación de algoritmos de alta precisión, como las Máquinas de Vectores de Soporte (SVM), lo que amplió sustancialmente el abanico de aplicaciones de la minería. En la actualidad, esta evolución ha derivado en la integración sinérgica de tecnologías de Big Data y Deep Learning, lo cual incrementa la capacidad de cómputo y afianza a la minería de datos como un elemento clave y dinámico de la ciencia de datos contemporánea, indispensable para obtener conocimiento estratégico a partir de conjuntos masivos y complejos.

3.1.2 Sistemas Informáticos

La integración de la ciencia de datos en la práctica historiográfica contemporánea ha provocado un desplazamiento ontológico del concepto de "dato", transitando de una unidad factual aislada hacia procesos complejos de datificación del pasado. Conforme a la investigación de Manrique-Gómez et al. (2025), este fenómeno no se limita a la digitalización de fuentes, sino que constituye una reconfiguración de la tríada tradicional documento-dato-hecho. En este sentido, los autores postulan que las bases de datos emergen como nuevas narrativas historiográficas no lineales. A la luz de lo expuesto por los investigadores, la historia digital actual emplea técnicas de *Machine Learning* para

gestionar el *Deep Data*: conjuntos de datos que permiten identificar patrones "invisibles" mediante lecturas distantes.

3.1.3 Evolución de la comunicación y la necesidad de análisis en el entorno digital

La forma de construir la realidad social mediante el lenguaje ha cambiado profundamente al integrar las Tecnologías de la Información y Comunicación (TIC) en la rutina diaria. Partiendo de la premisa de que es imposible no comunicar, la idea tradicional de "opinión pública" pasó de ser un modelo unidireccional dominado por medios clásicos a convertirse en un ecosistema mediático completamente interactivo de Dios et al. (2020).

3.1.4 Origen y Evolución de la Minería de Datos

La minería de datos se desprende del fenómeno del *big data* y sitúa sus orígenes históricos a finales de la década de 1940, fundamentándose inicialmente en la teoría de la información. Actualmente, estas técnicas permiten dar respuesta a la "explosión de información" mediante el descubrimiento de conocimiento (KDD) oculto en las bases de datos, permitiendo no solo describir realidades, sino también realizar modelados y pronósticos con fines predictivos sobre el comportamiento de los usuarios Valdiviezo et al. (2020)

3.1.5 Las Redes Sociales como Canales Institucionales: De la interacción social al compromiso organizacional

La transición de las plataformas digitales desde entornos orientados al ocio hacia infraestructuras estratégicas de comunicación ha redefinido la relación entre la universidad y su comunidad. Al articular dinámicas sociales con comunicación digital y relatos colaborativos, las instituciones superiores consiguen convertir la implicación individual en un recurso de gobernanza que fortalece la lealtad y la participación activa de los usuarios nativos digitales Galioto et al(2025).

3.2 Antecedente de investigación relacionados al tema presentado

3.2.1 Métricas de interacción y compromiso en redes sociales de universidades ecuatorianas

En el contexto nacional, investigaciones recientes han evaluado el impacto de las redes sociales en universidades públicas de la Zona 3 del Ecuador, destacando que la alta

penetración de internet en el país (81%) obliga a las instituciones a pasar de una presencia pasiva a una gestión estratégica de contenidos. Según D. E. M. Sánchez et al. (2025), el éxito de la comunicación institucional no reside únicamente en la cantidad de seguidores, sino en la capacidad de generar interacción a través de contenido cercano, emotivo y con sentido de pertenencia.

3.2.2 Algoritmos de predicción de engagement

Un antecedente fundamental para esta investigación es el estudio desarrollado por Chilon Vargas & Garibaldi Bravo (2025), quienes diseñaron un modelo predictivo basado en Machine Learning para determinar patrones de interacción en YouTube. Su metodología destaca por integrar técnicas de minería de texto aplicadas a los títulos y descripciones, junto con la extracción de metadatos mediante APIs oficiales para prever el éxito de un contenido antes de su publicación. Este enfoque demuestra que la combinación de características textuales y métricas de visualización permite anticipar el impacto de las publicaciones, proporcionando una base técnica sólida para el desarrollo de sistemas inteligentes enfocados en la optimización de la gestión de contenidos en redes sociales.

3.2.3 Minería de datos aplicada a la comunicación digital: Casos de éxito en otras instituciones educativas

Dentro del contexto nacional, destaca el estudio realizado por Domínguez-Gómez (2022), donde se aplicaron técnicas de *Educational Data Mining* (EDM) para analizar la percepción de la comunidad estudiantil sobre la comunicación y educación digital. Este escenario operativo deja al descubierto la falta de una infraestructura formal enfocada en monitorear los canales institucionales. Sin el respaldo de estas herramientas tecnológicas, el equipo encargado termina publicando a ciegas. Comparten el material digital sin tener cómo medir o evaluar de forma estructurada las respuestas de su audiencia. A la larga, operar sin el respaldo de indicadores exactos le resta peso a los mensajes institucionales y termina desgastando el posicionamiento general de la extensión académica.

3.2.4 Comportamiento del consumidor digital joven: Investigaciones sobre los hábitos de la Gen Z en Ecuador

Para comprender la dinámica de interacción de la muestra estudiada, es fundamental analizar los hallazgos de Serrano-Orellana et al. (2025). Al carecer de estas

herramientas tecnológicas, el equipo responsable termina publicando el contenido sin contar con un verdadero respaldo analítico. Difunden el material digital sin tener los medios necesarios para medir o evaluar de forma estructurada las reacciones de su audiencia. La investigación detalla que el 77 % de estos jóvenes accede a la red de forma exclusiva desde sus dispositivos móviles, consolidando al teléfono inteligente como el entorno principal para interactuar socialmente y consumir contenido.

3.3 Definiciones conceptuales

3.3.1.1 Sistemas Informáticos: El núcleo de la solución

Para el desarrollo de soluciones tecnológicas complejas, Terán et al. (2009) proponen un marco metodológico donde el sistema informático se define como una plataforma integral que debe ser evaluada bajo dominios tecnológicos específicos: seguridad, infraestructura, visualización y gestión de datos. Según este enfoque, el sistema no es solo una herramienta aislada, sino una arquitectura de software diseñada para alcanzar un desarrollo funcionalmente exitoso y plenamente integrado.

3.3.1.2 Arquitectura de Sistemas Web Modernos

La arquitectura de los sistemas web actuales ha evolucionado para resolver la problemática de la gestión heterogénea, donde diversas plataformas operan de manera independiente. Según Pérez Alvarez et al. (2024), la solución técnica más eficiente para este desafío es el empleo de Interfaces de Programación de Aplicaciones (APIs) como vehículos de comunicación fundamentales. En un modelo de arquitectura actual, el empleo de APIs REST posibilita que una aplicación única como la planteada en este proyecto incorpore información procedente de diversas fuentes incompatibles entre sí, estandarizando los tipos de datos en una base de datos centralizada.

3.3.1.3 Integración de Sistemas Heterogéneos mediante APIs

La administración de datos en ambientes tecnológicos diversos constituye uno de los retos más cruciales para la interoperabilidad de sistemas actuales. De acuerdo con Chilon Vargas & Garibaldi Bravo (2025), el uso de Interfaces de Programación de Aplicaciones (APIs) representa la alternativa técnica idónea para unificar información de distintos gestores en una plataforma única, solventando las incompatibilidades entre esquemas de integración tradicionales. En la ingeniería de software actual, las APIs no son

solo recursos tecnológicos, sino canales de comunicación esenciales que posibilitan la transferencia de información y la incorporación de funcionalidades a productos finales

3.3.2 Minería de Datos y el Proceso KDD

El descubrimiento de conocimiento en bases de datos (KDD) se define como un proceso estructurado y multietapa orientado a la identificación de patrones válidos, novedosos y potencialmente útiles dentro de conjuntos masivos de información. Este marco resulta fundamental para organizar el flujo sistemático de los datos desde su recolección inicial hasta la toma de decisiones estratégicas, mediante fases sucesivas de selección, preprocesamiento, transformación, minería de datos e interpretación o evaluación Pérez Álvarez et al.(2024)

3.3.3 Inteligencia Artificial y Machine Learning

La Inteligencia Artificial (IA) se define como una ciencia empírica orientada a simular procesos de cognición humana mediante métodos y modelos matemáticos. Dentro de este campo, el Aprendizaje Automático (*Machine Learning*) permite a los sistemas aprender de forma autónoma a partir de la experiencia y los datos, sin necesidad de una programación explícita para cada tarea. Según Villalobos-Murillo et al. (2025), la implementación exitosa de modelos predictivos requiere una sistematización que incluya la transformación de datos, el entrenamiento de redes neuronales y la validación mediante métricas de rendimiento como la precisión (*accuracy*) y la matriz de confusión.

3.3.3.1 Arquitectura de Sistemas Web Modernos e Interoperabilidad

La interoperabilidad tecnológica se concibe como la facultad de plataformas digitales diversas para transferir información de forma segura, eficaz y en tiempo real. Según Guzmán et al. (2025), el esquema de arquitecturas fundamentadas en APIs RESTful y microservicios constituye la alternativa idónea para articular aplicaciones móviles y plataformas web vigentes. Este modelo asegura que el tráfico de datos sea continuo y escalable, posibilitando una infraestructura adaptable a través de protocolos de autenticación sólidos.

3.3.4 Sistemas Informáticos: El núcleo de la solución

La transformación digital se define como un proceso evolutivo y estratégico que impacta a las organizaciones con el objetivo de expandir su presencia en el ecosistema digital mediante la innovación tecnológica. Según Calle Herencia (2022), la adopción de plataformas digitales permite a las instituciones encaminar procesos competitivos, optimizando sus recursos y ganando participación en su entorno operativo. En el marco de la ingeniería, un sistema informático no solo actúa como un almacén de datos, sino como el motor de este cambio, integrando pilares y componentes elementales que buscan la rentabilidad y la eficiencia.

3.3.5 Análisis de Sentimientos

En virtud de lo expresado por Benavides Morales & Herrera Flores (2022) , la especialista antes citada manifiesta que el análisis de sentimientos es una tarea del Procesamiento de Lenguaje Natural (PLN) centrada en la detección y categorización automática de las actitudes, opiniones y emociones expresadas en un texto. Este concepto se apoya en la utilización de algoritmos de clasificación que definen la polaridad de un segmento lingüístico, catalogándolo habitualmente como positiva, negativa o neutra. Teóricamente, el procedimiento exige la fragmentación del lenguaje no estructurado en vectores numéricos para que el sistema logre interpretar el peso subjetivo del mensaje, posibilitando la conversión de apreciaciones cualitativas en indicadores cuantitativos que muestran la condición emocional de un colectivo digital en un lapso temporal determinado.

3.3.5.1 Marketing de Contenidos

Atendiendo a los criterios de Estrada et al. (2021) , el marketing de contenidos es un proceso de gestión enfocado en identificar, anticipar y satisfacer los requerimientos del cliente de manera rentable, utilizando masivamente los medios digitales y las redes sociales. Este constructo se define como una experiencia interactiva fundamentada en la generación de valor percibido, donde las organizaciones crean y administran contenidos significativos basados en bits y distribuidos por canales electrónicos para atraer y retener consumidores. Conceptualmente, esta estrategia se articula a través de dimensiones esenciales como la instrucción, el entretenimiento y el interés, buscando diferenciar el

mensaje institucional en medio del ruido informativo para incidir positivamente en el comportamiento y el proceso de decisión de compra del segmento objetivo

3.3.5.2 Análisis Estructural

Bajo la óptica de Sarmiento et al. (2024), el análisis estructural se define como un examen académico riguroso que supera la apreciación superficial para descomponer la forma y dividir los componentes esenciales de un objeto específico. Este concepto se basa en la integración de principios técnicos y metodológicos que posibilitan un entendimiento minucioso de la disposición interna, favoreciendo la generación de un juicio crítico y un sustento conceptual sólido. En aspectos técnicos, este proceso facilita el establecimiento de un esquema analítico global donde se contemplan factores técnicos, estilísticos y expresivos, constituyéndose en la base para la lectura y evaluación científica de cualquier arquitectura o estructura compleja, haciendo que la comprensión del conjunto emane de la separación sistemática de sus partes.

3.3.6 Visualización de Datos (Data Visualization)

En correspondencia con el planteamiento de Mayet Comerón et al. (2022) , los académicos antes citados describen la visualización de datos como un proceso científico que posibilita la comprensión del sentido oculto de la información, el cual no resulta directamente aprehensible desde los datos, procesos o fenómenos aislados. Esta sección se fundamenta en la integración sinérgica de las cualidades comunicacionales de la imagen y del lenguaje verbal, orientada a desarrollar estrategias para la comunicación científica efectiva. A partir de este tratamiento, se estructuran especificaciones comprensibles que sirven como insumo para construir representaciones visuales manejables, aprovechando el procesamiento cognitivo natural del ser humano.

3.3.6.1 Tableros de Control Digital

Siguiendo la línea teórica de D. M. V. Sánchez et al. (2023) , el autor antes mencionado sostiene que el tablero de control o dashboard es una herramienta informática de visualización de datos diseñada para la representación gráfica de métricas y de indicadores clave de desempeño (KPI). La arquitectura de este sistema se sostiene en su capacidad para integrar técnicamente distintas fuentes de información. Al conectar estos flujos dispersos, logra proyectar una imagen unificada y al instante sobre el estado real de

un proceso o de toda una organización. Si lo analizamos desde una perspectiva más teórica, los cuadros de mando terminan funcionando como un pilar cognitivo indispensable para la inteligencia de negocios. En la práctica, su labor es bastante directa: toman bloques de datos pesados o difíciles de leer y los traducen en elementos visuales mucho más amigables. Es precisamente esta limpieza gráfica la que le permite al usuario detectar tendencias emergentes, fluctuaciones o nuevas ventanas de oportunidad sin tener que perderse en la complejidad de los números.

3.3.7 Inteligencia de Negocios

Tal como lo aseveran Arévalo Valarezo & Neira Picon (2023) , las autoras antes mencionadas definen la inteligencia de negocios (BI, por sus siglas en inglés) como un ecosistema tecnológico compuesto por un conjunto de aplicaciones y herramientas orientadas a la carga, análisis, extracción y reporte de información significativa. Este constructo se fundamenta en la capacidad de transformar flujos de datos brutos en activos de conocimiento que permiten orientar la toma de decisiones dentro de entornos complejos y competitivos.

3.3.8 Metodologías de Desarrollo de Software y el Modelo en Cascada (Waterfall)

Dentro de este dominio, el Modelo en Cascada (*Waterfall*) constituye el paradigma clásico del ciclo de vida secuencial, fundamentado en la ejecución lineal y progresiva de fases sucesivas e interdependientes: análisis de requisitos, diseño arquitectónico, implementación del código, pruebas de verificación y despliegue Pressman & Maxim (2020). Bajo este enfoque, cada etapa debe culminar y validarse formalmente antes de iniciar la subsiguiente, garantizando la trazabilidad documental, la definición rigurosa de entradas y salidas técnicas, y el control de calidad en cada componente del sistema informático.

La implementación del modelo en cascada resulta idónea en proyectos tecnológicos donde los objetivos del sistema, las fuentes de datos y las restricciones de integración se encuentran plenamente definidos y estabilizados desde las etapas preliminares de diagnóstico (Thesing et al., 2021). Esta linealidad permite establecer puntos de control rigurosos en cada fase del ciclo de vida, lo que facilita una documentación exhaustiva, la

trazabilidad de los flujos de datos y la mitigación temprana de fallos arquitectónicos o vulnerabilidades operativas antes del paso a producción Natarajan & Pichai (2024).

3.4 Conclusión

La articulación sistémica de los fundamentos teóricos presentados afianza un modelo donde la comunicación institucional supera la divulgación para transformarse en un procedimiento de ingeniería de datos de elevada exactitud. Mediante la convergencia entre el descubrimiento de conocimiento en bases de datos (KDD) y las estructuras web actuales basadas en APIs RESTful, se constituye una plataforma sólida capaz de atenuar la diversidad de las plataformas digitales vigentes, posibilitando una estandarización de datos que resulta crucial para la interoperabilidad. Esta arquitectura no solo actúa como un repositorio centralizado, sino como un motor de transformación donde el Procesamiento de Lenguaje Natural (PLN) y el Aprendizaje Automático específicamente mediante el entrenamiento de redes neuronales y la validación rigurosa a través de métricas de accuracy y matrices de confusión permiten la transmutación de variables cualitativas y percepciones emocionales en métricas de polaridad cuantitativa con fines predictivos. Gestionar la comunicación de forma tan empírica revela una desconexión profunda y evidencia una brecha significativa entre la información que se difunde y el contenido que la comunidad realmente desea consumir. Dado que el teléfono inteligente se ha consolidado indiscutiblemente como su principal canal de consumo, cualquier estrategia de difusión debe adaptarse a estos patrones específicos. Es precisamente aquí donde la integración tecnológica marca la diferencia.

Fusionar la inteligencia de negocios con una visualización limpia, apoyada en cuadros de mando dinámicos, entrega a la dirección una herramienta estratégica de primer nivel para justificar cada una de sus acciones. A fin de cuentas, la verdadera utilidad del sistema radica en su capacidad operativa: brindar el poder de auditar, evaluar y mantener bajo control el impacto de la comunicación institucional en el momento exacto en que ocurre.

CAPÍTULO III

4 MARCO INVESTIGATIVO

4.1 Introducción

El presente apartado constituye la estructura medular de la investigación, donde se definen los procedimientos, métodos y técnicas sistemáticas que permiten la transición del constructo teórico hacia la implementación tecnológica. Para el desarrollo del "Sistema Informático con Minería de Datos para la difusión de contenido en las redes sociales de la ULEAM El Carmen", el diseño metodológico no se limita a una simple descripción de pasos, sino que se establece como el protocolo de rigor científico que valida la recolección, el procesamiento y la interpretación de los activos digitales institucionales.

La índole de esta propuesta demanda un marco investigativo multidimensional. Por un lado, requiere el rigor de la investigación aplicada para trasladar los conceptos de la ciencia de datos a un aplicativo funcional; por otro, se sustenta en un enfoque cuantitativo con alcance correlacional para analizar con objetividad las pautas de interacción en plataformas como Facebook, Instagram y YouTube.

Este esquema se sustenta en la premisa de que la efectividad en la propagación de contenidos puede ser cuantificada y perfeccionada mediante la evidencia empírica. Por tanto, el marco que se detalla a continuación articula la lógica de la ingeniería de software con el rigor de la metodología científica, asegurando que cada algoritmo implementado y cada métrica analizada responda a una necesidad real del equipo de Softmedia y de la comunidad universitaria, garantizando la replicabilidad y la validez técnica de los hallazgos obtenidos.

4.2 Tipo de investigación

4.2.1 Investigación Aplicada

Según lo planteado por Duoc (2025), la investigación aplicada consiste en trabajos originales ejecutados para adquirir conocimientos dirigidos hacia un objetivo práctico específico. Este constructo busca determinar métodos para alcanzar metas predeterminadas mediante la aplicación de teorías científicas, permitiendo solucionar problemas concretos

dentro de un sector social o productivo a través de la implementación de resultados tangibles.

En el presente estudio, este tipo de investigación cobró forma mediante el diseño, desarrollo y despliegue del sistema informático con minería de datos para el equipo SoftMedia de la ULEAM, Extensión El Carmen. Se articularon fundamentos de la ingeniería de software, minería de datos y procesamiento de lenguaje natural (PLN) para dar solución al problema de dispersión de métricas, ausencia de análisis de percepción y gestión manual de contenidos en las redes institucionales.

4.2.2 Investigación Descriptiva

En la opinión de Guevara Albán et al. (2020) , la investigación descriptiva se orienta a detallar características fundamentales de fenómenos homogéneos empleando criterios sistemáticos. Este proceso permite establecer la estructura y el comportamiento del objeto de estudio, generando información técnica comparable que facilita la delimitación del problema. Es esencial para identificar patrones iniciales antes de proceder con el análisis profundo de los datos recolectados.

Se orientó a caracterizar el perfil demográfico, los hábitos de navegación y las preferencias de consumo digital de la comunidad universitaria en la ULEAM, Extensión El Carmen. A partir del instrumento administrado a una muestra de 2.074 personas, este enfoque permitió relevar y describir variables clave como los rangos etarios predominantes, la demanda de formatos audiovisuales multimodales y las franjas horarias de mayor interacción en las plataformas oficiales.

4.2.3 Investigación Correlacional

En su investigación Zuasnábar et al. (2023), definen la investigación correlacional como un método cuantitativo de diseño no experimental que busca establecer el grado de relación estadística entre dos o más variables. Su propósito central es medir y comprender la conexión entre dichas dimensiones, lo que permite predecir comportamientos al evaluar la frecuencia y la dirección de las asociaciones halladas en una muestra específica.

Este enfoque permitió examinar la relación existente entre las características de los contenidos difundidos (formato audiovisual, duración, temática y franja horaria de

publicación) y las métricas de respuesta generadas por la audiencia (*engagement*, número de reproducciones y polaridad emocional de los comentarios) en las diferentes redes sociales institucionales, facilitando la parametrización de las reglas analíticas del sistema informático.

4.3 Métodos de investigación

4.3.1 Método Cuantitativo

De acuerdo con Flores & Zamora (2023), el método cuantitativo se apoya en el paradigma positivista para recopilar y analizar datos numéricos. Este enfoque emplea la lógica deductiva para comprobar teorías y evaluar hipótesis mediante procedimientos estandarizados y herramientas estadísticas. Este método se canalizó en la recolección, sistematización y análisis numérico de los datos levantados mediante el instrumento diagnóstico aplicado a la comunidad universitaria. Asimismo, se utilizó en la medición cuantitativa de las variables técnicas de rendimiento del sistema, tales como los tiempos de extracción de datos a través de APIs oficiales, el volumen de registros almacenados en la base de datos y la velocidad de respuesta en la generación de reportes automatizados.

4.3.2 Método Descriptivo

En la opinión de Guevara Albán et al. (2020), la investigación descriptiva busca explicar las características principales de los fenómenos bajo criterios sistemáticos. Esta metodología ayuda a determinar la estructura o el comportamiento del tema estudiado, generando información contrastable con otras fuentes. Es clave para reconocer patrones y rasgos particulares de un grupo homogéneo, lo cual permite entender a fondo la situación observada antes de ejecutar cualquier intervención técnica.

Se recurrió a este método para caracterizar de forma detallada y ordenada el flujo de trabajo actual del equipo SoftMedia, identificando los hábitos de consumo digital de la comunidad universitaria y las limitantes operativas asociadas a la gestión manual de métricas en plataformas como Facebook, Instagram, TikTok y YouTube.

4.3.3 Método Analítico

Desde la óptica de Vergara-Gonzalez et al. (2020), el método analítico se define como un procedimiento sistemático orientado a descomponer procesos complejos en sus

componentes elementales para optimizar la planificación operativa. Esta metodología permite identificar variables críticas y causas de no conformidad, facilitando la toma de decisiones basada en datos precisos y la mejora continua de los resultados mediante el análisis pormenorizado de cada elemento constituyente.

Este método se materializó en la descomposición técnica del sistema global en sus módulos constituyentes: la ingesta automatizada de datos desde APIs externas, el procesamiento de lenguaje natural (PLN) para la clasificación de sentimientos en comentarios, la persistencia relacional en PostgreSQL bajo Docker y la renderización en el panel de control (*dashboard*) desarrollado en React, permitiendo validar y optimizar cada capa de la arquitectura de forma independiente.

4.4 Fuentes de información

4.4.1 Encuestas Online

Bajo el análisis de Torrejón-Guirado et al. (2024) , las encuestas online se definen como instrumentos digitales de recolección de datos que requieren una rigurosa optimización técnica y metodológica. Este constructo integra protocolos de ética digital, validación de diseño y procesos sistemáticos de reclutamiento virtual, facilitando el manejo y análisis de grandes volúmenes de información para medir comportamientos y percepciones dentro de entornos de alta interacción digital.

Esta técnica se instrumentó mediante el diseño y despliegue de un cuestionario estructurado de 10 ítems a través de la plataforma institucional Microsoft Forms. La herramienta permitió canalizar la recolección de datos de manera automatizada entre los estamentos de la comunidad universitaria (estudiantes, docentes y personal administrativo) , garantizando la autenticación de los participantes, el registro cronológico exacto de inicio y finalización de cada respuesta , y la posterior exportación estructurada del dataset hacia hojas de cálculo para su respectiva tabulación, análisis estadístico y parametrización de los requerimientos del sistema SoftMedia.

4.5 Estrategia operacional para la recolección de datos

4.5.1 Población

Según Manterola et al. (2024), la población se define como el conjunto total de elementos o individuos que comparten características comunes y constituyen el escenario de estudio. Este componente es determinante en el proceso investigativo, ya que delimita el alcance de la recolección de datos y condiciona la metodología, permitiendo que los resultados obtenidos posean un nivel de evidencia sólido y representativo.

Para el desarrollo de la presente investigación, se ha identificado una población total de 2627 participantes, vinculados directamente con la planta principal de la ULEAM El Carmen. Dicho colectivo se encuentra distribuido entre el cuerpo estudiantil, el profesorado y el personal administrativo, garantizando una visión integral de los usuarios del sistema.

4.5.1.1 Segmentación

Personal administrativo, profesores y estudiantes

4.5.1.2 Técnica de muestreo

Para determinar el número de unidades de análisis necesarias, se utiliza una fórmula que integra el tamaño de la población total (N), el grado de seguridad de los resultados (Z), la probabilidad de ocurrencia del evento (p) y el error de estimación permitido (e). En el caso de poblaciones finitas, este procedimiento permite ajustar el tamaño de la muestra de manera proporcional al total de individuos conocidos, tal como se detalla a continuación:

$$n = \frac{N * z^2 * p * (1 - p)}{(N - 1) * e^2 + z^2 * p * (1 - p)}$$

Para la determinación del tamaño muestral, se han considerado los siguientes parámetros:

N = 2497 (Tu población total)

Z = 1.96 (Para un nivel de confianza del 95%)

p = 0.5 (Probabilidad de éxito)

q = 0.5 (Probabilidad de fracaso, 1 - p)

$e = 0.05$ (Margen de error del 5%)

Una vez definidos los criterios de confianza y error, se procede a la sustitución de las variables en la fórmula para determinar el tamaño de la muestra final:

$$n = \frac{2497 * (1.96)^2 * 0.5 * 0.5}{(0.05)^2 * (2497 - 1) + (1.96)^2 * 0.5 * 0.5}$$

4.5.1.3 Tamaño de la muestra

Considerando una población finita de 2607 individuos, se determinó mediante la fórmula de tamaño muestral que se requieren 334 encuestas para garantizar una representatividad estadística. Este cálculo se basa en un margen de error aceptable del 5% y un coeficiente de confianza del 95%.

4.5.2 Análisis de herramientas de recolección

4.5.2.1 Encuesta

Encuesta dirigida a: Estamentos de la ULEAM El Carmen (Planta Central): cuerpo estudiantil, personal académico y administrativo.

Propósito: Evaluar los requerimientos de usuario y preferencias de consumo digital para el desarrollo de un sistema informático especializado en la gestión y propagación de contenidos digitales en las plataformas sociales de la institución.

Institución: Universidad Laica Eloy Alfaro de Manabí Extensión El Carmen.

4.5.3 Plan de recolección de datos

Tabla 1: Recolección de datos

Elemento	Descripción Técnica
Población	2,074 individuos (Estudiantes, Docentes, Administrativos).
Lugar	Planta Central - ULEAM Extensión El Carmen.
Técnica	Encuesta.
Instrumento	Cuestionario estructurado de 10 ítems.
Medio	Digital / Plataforma de formularios.
Propósito	Captura de metadatos para el sistema de minería de datos

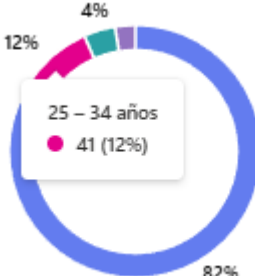
Tiempo	25 días
--------	---------

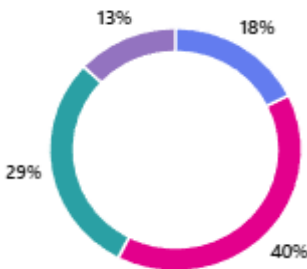
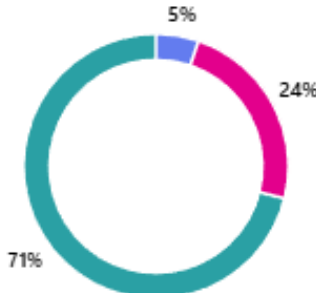
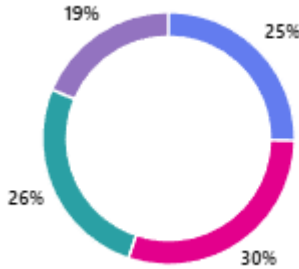
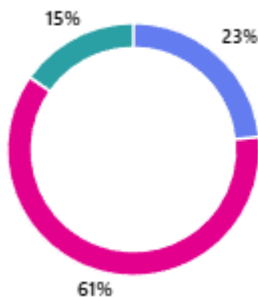
4.6 Análisis de resultados

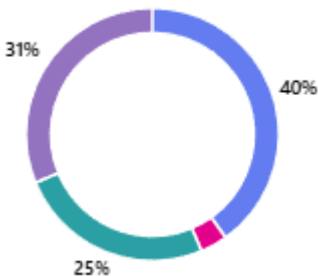
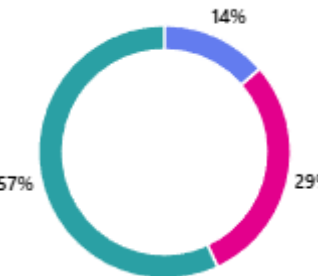
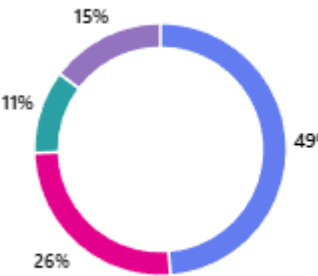
4.6.1 Tabulación y análisis de los datos

4.6.1.1 Aplicación de encuestas a docentes, estudiantes y personal administrativo de la ULEAM, extensión El Carmen, en el área de la planta central.

Tabla 2: Tabulación y análisis de datos de encuesta

Preguntas	Respuestas	Intervención
¿Cuál es su rol principal dentro de la planta central ULEAM El Carmen?	 96%	El 96 % de estudiantes define al alumnado como el usuario objetivo principal. El sistema orienta los algoritmos de minería de datos y Procesamiento de Lenguaje Natural (PLN) hacia las opiniones e interacciones del colectivo estudiantil
2. Indique su rango de edad	 82% 25 – 34 años 41 (12%)	El 82 % concentrado entre 18 y 24 años ratifica un perfil nativo digital. Esto exige interfaces web dinámicas en React y la extracción prioritaria de redes sociales de consumo rápido.
3. ¿Qué orientación visual garantiza una experiencia de usuario (UX) más satisfactoria al consumir videos institucionales?	 59% 41%	La división entre formato horizontal (59 %) y vertical (41 %) requiere que el sistema registre y clasifique las métricas según la relación de aspecto del video para evaluar el impacto en móviles frente a ordenadores.

Preguntas	Respuestas	Intervención										
4. ¿Cuál es el rango de tiempo máximo que usted está dispuesto a invertir en un video antes de que su atención decaiga?	 <table><caption>Data for Question 4: Video Length Preference</caption><thead><tr><th>Color</th><th>Percentage</th></tr></thead><tbody><tr><td>Pink</td><td>40%</td></tr><tr><td>Teal</td><td>29%</td></tr><tr><td>Blue</td><td>18%</td></tr><tr><td>Purple</td><td>13%</td></tr></tbody></table>	Color	Percentage	Pink	40%	Teal	29%	Blue	18%	Purple	13%	La preferencia por videos de 20 a 60 segundos (40 %) establece el umbral de retención para parametrizar las reglas de recomendación del sistema y alertar sobre contenidos extensos.
Color	Percentage											
Pink	40%											
Teal	29%											
Blue	18%											
Purple	13%											
5. Al ver un video en redes sociales, ¿Qué importancia les da a los subtítulos?	 <table><caption>Data for Question 5: Subtitle Importance</caption><thead><tr><th>Color</th><th>Percentage</th></tr></thead><tbody><tr><td>Teal</td><td>71%</td></tr><tr><td>Pink</td><td>24%</td></tr><tr><td>Blue</td><td>5%</td></tr></tbody></table>	Color	Percentage	Teal	71%	Pink	24%	Blue	5%	El 71 % demanda videos con audio y subtítulos. El sistema contempla la auditoría de soporte textual en las publicaciones para optimizar la minería de texto y la comprensión del mensaje.		
Color	Percentage											
Teal	71%											
Pink	24%											
Blue	5%											
6. ¿Qué tipo de información le gustaría recibir a través de videos?	 <table><caption>Data for Question 6: Information Type Preference</caption><thead><tr><th>Color</th><th>Percentage</th></tr></thead><tbody><tr><td>Pink</td><td>30%</td></tr><tr><td>Teal</td><td>26%</td></tr><tr><td>Blue</td><td>25%</td></tr><tr><td>Purple</td><td>19%</td></tr></tbody></table>	Color	Percentage	Pink	30%	Teal	26%	Blue	25%	Purple	19%	El interés centrado en eventos institucionales (30 %) y trámites académicos (25 %) define las categorías temáticas en la base de datos PostgreSQL para medir el rendimiento por tipo de contenido.
Color	Percentage											
Pink	30%											
Teal	26%											
Blue	25%											
Purple	19%											
7. ¿Qué tono prefiere que tengan los videos de la universidad?	 <table><caption>Data for Question 7: Video Tone Preference</caption><thead><tr><th>Color</th><th>Percentage</th></tr></thead><tbody><tr><td>Pink</td><td>61%</td></tr><tr><td>Blue</td><td>23%</td></tr><tr><td>Teal</td><td>15%</td></tr></tbody></table>	Color	Percentage	Pink	61%	Blue	23%	Teal	15%	La preferencia por un tono dinámico y juvenil (61 %) condiciona el modelo de análisis de sentimientos para que reconozca modismos y lenguaje informal dentro de los comentarios.		
Color	Percentage											
Pink	61%											
Blue	23%											
Teal	15%											

Preguntas	Respuestas	Intervención										
8. ¿En qué plataforma consume más contenido relacionado con la ULEAM?	 <table><caption>Platform Usage Data</caption><thead><tr><th>Platform</th><th>Percentage</th></tr></thead><tbody><tr><td>Facebook</td><td>40%</td></tr><tr><td>TikTok</td><td>31%</td></tr><tr><td>YouTube</td><td>25%</td></tr><tr><td>Instagram</td><td>4%</td></tr></tbody></table>	Platform	Percentage	Facebook	40%	TikTok	31%	YouTube	25%	Instagram	4%	El uso de Facebook (40 %) junto a TikTok, Instagram y YouTube sustenta la arquitectura desacoplada del sistema, integrando mineros independientes para cada API oficial.
Platform	Percentage											
Facebook	40%											
TikTok	31%											
YouTube	25%											
Instagram	4%											
9. ¿En qué horario suele revisar con más frecuencia sus redes sociales?	 <table><caption>Social Media Usage by Time</caption><thead><tr><th>Time</th><th>Percentage</th></tr></thead><tbody><tr><td>Nocturnal (18:00 a 23:00)</td><td>57%</td></tr><tr><td>Daytime</td><td>29%</td></tr><tr><td>Morning</td><td>14%</td></tr></tbody></table>	Time	Percentage	Nocturnal (18:00 a 23:00)	57%	Daytime	29%	Morning	14%	El 57 % de concurrencia nocturna (18:00 a 23:00) programa la ejecución del orquestador automatizado y define las sugerencias de publicación en horas de máximo tráfico.		
Time	Percentage											
Nocturnal (18:00 a 23:00)	57%											
Daytime	29%											
Morning	14%											
10. ¿A través de qué formato prefiere recibir la información oficial de la universidad?	 <table><caption>Preferred Content Format Data</caption><thead><tr><th>Format</th><th>Percentage</th></tr></thead><tbody><tr><td>Video</td><td>49%</td></tr><tr><td>Static Images</td><td>26%</td></tr><tr><td>Audio</td><td>15%</td></tr><tr><td>Text</td><td>11%</td></tr></tbody></table>	Format	Percentage	Video	49%	Static Images	26%	Audio	15%	Text	11%	El 49 % prefiere micro contenido audiovisual (Reels y tutoriales). El dashboard prioriza la visualización y auditoría de métricas de interacción en formatos de video frente a imágenes estáticas.
Format	Percentage											
Video	49%											
Static Images	26%											
Audio	15%											
Text	11%											

4.6.2 Presentación y descripción de los resultados obtenido

El análisis de los instrumentos aplicados al 33.3 % de los 2,074 integrantes de la planta central muestra una postura favorable hacia la implementación de una infraestructura informática que optimice la comunicación institucional. Un dato determinante de la caracterización inicial es que los estudiantes conforman el 96 %, lo que define el perfil de usuario principal para entrenar los algoritmos de minería de datos.

Los resultados de la investigación también indican que los jóvenes de 18 a 24 años representan el 82 % de la población total. Este hallazgo confirma que el sistema debe ajustarse a patrones de consumo digital muy dinámicos. En cuanto al aspecto visual, existe una clara preferencia por el formato vertical, una tendencia que define la arquitectura de la solución y exige priorizar los dispositivos móviles para asegurar una óptima experiencia de usuario (UX).

Así, cada componente técnico responde con precisión a las costumbres reales de la audiencia.

Si analizamos la efectividad del mensaje, los hallazgos del cuarto ítem del estudio revelan una barrera de tiempo bastante estricta. Básicamente, el 40% de los encuestados solo logra mantener la atención en una ventana muy corta, la cual oscila entre los 20 y los 60 segundos. Este comportamiento no es un hecho aislado, sino que conecta de inmediato con los datos que arrojó la quinta pregunta. Ahí quedó confirmada la fuerza que tiene el consumo multimodal dentro de la comunidad. Para ponerlo en cifras, un contundente 71% de la muestra señaló que prefiere ver los videos activando el audio y los subtítulos al mismo tiempo. Encontrar una tendencia tan marcada nos deja una lección clara sobre cómo diseñar los materiales digitales.

Por otro lado, al evaluar la jerarquización de los temas, los datos de la sexta pregunta marcan una tendencia indiscutible. En la práctica, el mayor nivel de demanda se concentra en los procesos de matriculación y la gestión de trámites académicos. Si tomamos esta necesidad informativa y la combinamos con el uso dominante de Facebook y la preferencia de los usuarios por un tono mucho más dinámico, el sistema de minería de datos encuentra exactamente el escenario que necesita. Para completar el panorama, los hallazgos correspondientes a las preguntas nueve y diez terminan de comprobar la viabilidad del proyecto. Saber que más del 49% de la comunidad universitaria respalda la idea de recibir los comunicados oficiales mediante formatos audiovisuales, justamente durante las franjas horarias de mayor conexión, funciona como una sólida garantía. Es este nivel de aceptación el que, en definitiva, asegura que el material de la institución logre propagarse de manera orgánica y alcance un impacto verdaderamente positivo.

4.6.3 Informe final del análisis de los datos.

El procesamiento del dataset recolectado, correspondiente al 33.3% de la población objetivo (2,074 individuos), permite validar la necesidad tecnológica de un sistema de difusión automatizado en la ULEAM El Carmen. El análisis exhaustivo de las variables métricas y psicográficas arroja los siguientes pilares de diseño para el sistema informático:

Identificación del Clúster Primario: La muestra revela una homogeneidad crítica con un 96% de participación estudiantil y un 81% de usuarios en el rango etario de 18 a 24 años. Estos datos configuran un perfil de usuario nativo digital, lo que exige que el sistema de minería de datos procese contenidos de alta rotación y adaptabilidad estética.

Determinación de Parámetros UX/UI: La consolidación de la tendencia hacia el formato vertical define la arquitectura técnica de la interfaz, la cual debe estar orientada exclusivamente al despliegue en terminales móviles para asegurar la compatibilidad con el hábito de consumo detectado.

Heurísticas de Retención y Comprensión: Se determina un margen de atención acotado, donde el 39% del público privilegia contenidos de breve duración (20-60 segundos). Asimismo, la demanda de multimodalidad resulta manifiesta; el 71% de la población requiere la articulación de audio y soporte textual sincronizado, consolidando la redundancia informativa como una condición funcional indispensable para el éxito de la difusión.

Las evidencias obtenidas confirman la factibilidad del proyecto, puesto que más del 48% del colectivo universitario no solo aprueba la utilización de formatos audiovisuales, sino que exige una fuente de información oficial tecnificada. El sistema informático propuesto se posiciona como una herramienta de alto impacto para la propagación orgánica, garantizando que el flujo de datos institucionales sea eficiente, segmentado y coherente con el comportamiento digital de la planta central.

CAPÍTULO IV

5 MARCO PROPOSITIVO (ELABORACIÓN DE LA PROPUESTA)

5.1 Introducción

Este apartado detalla la arquitectura y el despliegue técnico del sistema inteligente diseñado para la analítica y predicción del impacto digital en las plataformas sociales del equipo de Softmedia, dentro de la ULEAM, Extensión El Carmen. La propuesta enlazar una gestión estratégica de los recursos humanos, tecnológicos y financieros, trazando un camino claro para transformar datos primarios en conocimiento accionable. Con el propósito de garantizar la viabilidad del aplicativo, se realiza una especificación detallada de sus requerimientos funcionales y no funcionales.

En lo que respecta a la gestión del ciclo de vida del proyecto, se adopta un modelo de desarrollo en cascada (Waterfall). Esta metodología garantiza un flujo estructurado y secuencial, permitiendo que cada fase desde el levantamiento inicial de requisitos y el modelado de interfaces, hasta la implementación del código y las pruebas de aseguramiento de calidad se ejecute bajo un control riguroso de integridad y análisis de resultados. Este enfoque lineal facilita la validación de cada componente antes de avanzar a la etapa de despliegue, asegurando que el sistema responda fielmente a las necesidades de auditoría digital de la institución.

La ejecución tecnológica se fundamenta en un conjunto de herramientas modernas y de elevada eficiencia. En cuanto a la lógica de servidor (*backend*), se emplean Node.js y Express.js, en tanto que la interfaz de usuario (*frontend*) se construye mediante React y Tailwind CSS, optimizando la presentación de indicadores y la experiencia del usuario. La administración del almacenamiento relacional se realiza con Postgres:15, mientras que la seguridad en los accesos se respalda a través del estándar JSON Web Tokens (JWT). Por otro lado, el motor de Inteligencia Artificial se programa íntegramente en Python, apoyándose en herramientas líderes de la industria como Figma para el diseño de interfaces, Postman para la documentación de la API y GitHub como eje central para el control de versiones y la colaboración continua.

5.2 Descripción de la propuesta

La propuesta consiste en el diseño, desarrollo y despliegue de una plataforma informática sustentada en técnicas de minería de datos y procesamiento de lenguaje natural (PLN), orientada a auditar, analizar y predecir el impacto de los contenidos digitales oficiales gestionados por el equipo SoftMedia de la ULEAM, Extensión El Carmen.

Para garantizar un ciclo de vida de desarrollo ordenado, trazable y riguroso, la propuesta se estructura bajo la metodología del Modelo en Cascada (*Waterfall*), ejecutando de manera secuencial cuatro fases técnicas fundamentales:

- **Fase I. Análisis de requisitos:** Levantamiento, depuración y especificación técnica de los requerimientos funcionales y no funcionales a partir del diagnóstico de hábitos de consumo digital de la comunidad universitaria.
- **Fase II. Diseño del sistema:** Modelado estructural mediante diagramas UML (casos de uso, secuencia y componentes), diseño físico del esquema relacional en PostgreSQL y maquetación de las vistas del panel de control (UI/UX).
- **Fase III. Desarrollo e implementación:** Codificación modular de la capa de servicios con FastAPI (Python), construcción de la interfaz dinámica en React con Tailwind CSS, y despliegue de los mineros autónomos para la extracción de métricas desde YouTube, TikTok, Facebook e Instagram.
- **Fase IV. Pruebas y verificación de software:** Ejecución sistemática de pruebas de caja blanca, caja negra y de integración del *pipeline* de datos para certificar la estabilidad de la persistencia y la exactitud del análisis de polaridad.

Como resultado final, la plataforma entrega un panel de control interactivo que visualiza tendencias críticas en tiempo real, facilitando la toma de decisiones respaldada en analítica de datos y optimizando la efectividad de la comunicación institucional.

5.3 Determinación de recursos

5.3.1 Humanos

La construcción y el funcionamiento del sistema requieren la participación de actores estratégicos, clasificados según su grado de implicación técnica y administrativa en el proyecto.

Tabla 3: Roles y funciones de los actores estratégicos del sistema

Personal	Función
Programador	Responsable del diseño arquitectónico, la implementación del motor de IA en Python y el desarrollo de la interfaz en React.
Equipo Softmedia	Gestión de contenidos y análisis estratégico del impacto digital en redes sociales.
Administrador del Sistema	Encargado de supervisar la persistencia de la información en Postgres:15, auditar el flujo de datos y validar la precisión de los reportes generados por el sistema.

5.3.2 Tecnológicos

La implementación del sistema demanda una infraestructura tecnológica que articule capacidad de cómputo para los modelos de IA y solidez en el despliegue de los servicios web, asegurando un equilibrio entre la eficiencia operativa y la factibilidad económica.

Tabla 4: Recurso humano

Personal	Función
Programador	Responsable del diseño arquitectónico, la implementación del motor de IA en Python y el desarrollo de la interfaz en React .
Equipo Softmedia	Gestión de contenidos y análisis estratégico del impacto digital en redes sociales.
Administrador del Sistema	Encargado de supervisar la persistencia de la información en Postgres:15, auditar el flujo de datos y validar la precisión de los reportes generados por el sistema.

Tabla 5: Recursos de hardware

Hardware	Especificaciones
Computadora de desarrollo	Procesador Intel core i5 o superior, Memoria RAM de 8GB o superior y Almacenamiento SSD de 240GB.
Servidor	HP ProLiant ML350 G6 Server o infraestructura con capacidad de procesamiento para entornos Node.js y Python.

Tabla 6: Recursos de software y especificaciones.

Categoría	Herramienta
Backend	Python, FastAPI y servidor ASGI Uvicorn.
Frontend	React y Tailwind CSS.
Base de Datos	PostgreSQL y SQLAlchemy.
Extracción de Datos	APIs oficiales de YouTube y TikTok, y subprocesos automatizados.
Inteligencia Artificial	Python
Diseño y Control	Figma, Postman

Económico

Los recursos económicos representan la inversión financiera necesaria para la adquisición de suministros, servicios y materiales de oficina indispensables para el desarrollo, pruebas y documentación final del proyecto integrador.

Tabla 7: Recursos económicos.

Rubro	Descripción	Costo estimado (USD)
Servicio de Internet	Conectividad mensual para acceso a librerías, GitHub y despliegue.	\$120.00
Materiales de oficina	Resmas de papel, tóner y útiles necesarios para la documentación.	\$40.00
Suministros de impresión	Impresión y encuadernación de los ejemplares de la tesis para titulación.	\$60.00
Energía eléctrica	Costo operativo por el uso de la estación de trabajo y el servidor.	\$50.00
Total	Inversión estimada para la ejecución del proyecto	\$270.00

5.4 Desarrollo

La puesta en marcha de la propuesta tecnológica se sustenta en la articulación sistémica de tres ejes fundamentales: el factor humano, tecnológico y financiero, los cuales aseguran la viabilidad operativa y la sostenibilidad de la iniciativa. El talento humano representa el núcleo ejecutor, con la participación de los desarrolladores en la implementación técnica y del equipo de Softmedia como beneficiario directo y responsable de la lectura estratégica de los datos.

5.4.1.1 Fase I Análisis de requisitos

Dentro del esquema de la metodología en cascada adoptado para esta iniciativa, la etapa inicial de análisis representa el pilar operativo esencial para la construcción del sistema de auditoría digital de Softmedia. Este flujo contempla la obtención, depuración y ratificación técnica de las necesidades del software, clasificadas rigurosamente en

requerimientos funcionales y no funcionales. Dicha división facilita la alineación entre las facultades analíticas del motor de IA concebido en Python y la interfaz reactiva en React con las metas estratégicas de la entidad, garantizando una arquitectura de datos consistente, eficiente y escalable.

5.4.1.2 Requisitos Funcionales

Las exigencias funcionales constituyen las demandas directas sobre el funcionamiento del sistema, delimitando con exactitud las tareas del software y su modo de respuesta frente a las solicitudes de los usuarios. En el marco de la solución concebida para Softmedia, estas se establecieron a través de un examen exhaustivo de la problemática existente y el levantamiento de información proveniente del personal técnico y del colectivo universitario de la ULEAM, Extensión El Carmen. A partir de este diagnóstico, se establecieron las funcionalidades críticas que el sistema debe ejecutar para garantizar la captura, el procesamiento mediante IA y la visualización del impacto digital en redes sociales:

- **Gestión de Autenticación y Perfiles:** El sistema debe permitir el inicio de sesión y cierre de sesión de los usuarios mediante credenciales seguras, utilizando JSON Web Tokens (JWT) para la gestión de sesiones.
- **Registro de Publicaciones:** El software debe permitir el ingreso y almacenamiento de los datos de las publicaciones de redes sociales (texto, métricas iniciales, fecha) en la base de datos Postgres:15.
- **Minería de Datos Automatizada:** La plataforma debe incorporar componentes de extracción autónomos que se ejecuten en segundo plano para capturar métricas de interacción, reproducciones y comentarios recientes mediante las APIs oficiales de plataformas como YouTube y TikTok.
- **Análisis de Sentimientos:** Para analizar los textos publicados, la plataforma utiliza un motor de inteligencia artificial programado en Python. En la práctica, esto permite medir con suma exactitud la polaridad emocional de la audiencia, clasificando su reacción frente al contenido como positiva, negativa o neutra.
- **Predicción de Impacto:** El motor de inteligencia artificial debe predecir el éxito y el alcance potencial de cada publicación. Para alcanzar este objetivo, fundamenta

sus proyecciones en el análisis de patrones históricos extraídos a partir de la información recopilada.

- **Visualización de Métricas:** La interfaz, desarrollada en React, integrará cuadros de mando y gráficos dinámicos, lo que permitirá que el sistema refleje en tiempo real el nivel de interacción de la audiencia y las tendencias generadas por cada publicación.
- **Generación de Reportes:** La plataforma debe permitir la exportación de informes detallados sobre el rendimiento de los contenidos, lo que en la práctica proporciona un respaldo técnico sólido y confiable, agilizando los procesos de evaluación y auditoría a cargo del equipo de Softmedia.

5.4.1.3 Requisitos no Funcionales

Los requerimientos no funcionales establecen los atributos de calidad, las restricciones y los parámetros críticos que la plataforma de Softmedia debe cumplir para asegurar un rendimiento óptimo, seguro y eficiente. Estas exigencias no describen las operaciones específicas del software, sino que determinan cómo debe comportarse el sistema bajo condiciones particulares de operación y seguridad.

- **Integración del Frontend y Backend:** Conexión eficiente y de baja latencia entre la interfaz desarrollada en React y el servidor en **FastAPI (Python)** mediante una API RESTful.
- **Optimización de Consultas:** Garantizar una respuesta rápida (preferiblemente menor a 2 segundos) en el filtrado de los registros de impacto digital y en la carga de métricas del *Dashboard*.
- **Interoperabilidad del Motor de IA:** Conexión fluida y asíncrona entre el servidor de aplicaciones y los submódulos de Python encargados del procesamiento inteligente de datos y la minería automatizada.
- **Rendimiento del Análisis de Sentimientos:** Capacidad de procesamiento eficiente por parte de la IA para la categorización masiva de emociones (positivas, negativas y neutras) sin saturar los recursos de memoria del servidor.

- **Persistencia y escalabilidad de datos:** El diseño, la estructuración y el almacenamiento en una base de datos relacional con PostgreSQL y SQLAlchemy aseguran que el sistema escale conforme aumenta el volumen de publicaciones procesadas.
- **Seguridad en el flujo de datos:** Establecer un canal de comunicación cifrado entre la base de datos y el backend protege la información en tránsito, lo que garantiza la confidencialidad del material institucional crítico y previene filtraciones no autorizadas.
- **Protección de rutas:** Se aplicarán controles de acceso y esquemas de autenticación rigurosos mediante el estándar JSON Web Tokens (JWT), con el fin de restringir el ingreso a los módulos administrativos exclusivamente al personal autorizado.
- **Integridad de la Base de Datos:** La implementación de protocolos de seguridad y controles rigurosos de acceso al motor PostgreSQL previene directamente las vulnerabilidades por inyección SQL. Con esta medida, se bloquea cualquier alteración no autorizada sobre la estructura o el contenido de la información.

5.4.1.4 Requisitos de software y hardware

A fin de garantizar un despliegue efectivo y un rendimiento óptimo de la plataforma concebida para Softmedia, resulta indispensable cumplir con un conjunto de especificaciones técnicas mínimas capaces de respaldar tanto el procesamiento de Inteligencia Artificial como la carga operativa de la interfaz web.

- **Servidor de Aplicaciones:** Infraestructura robusta con capacidad para alojar el entorno de ejecución Node.js y el motor de base de datos Postgres:15; se sugiere el uso de servidores físicos tipo HP ProLiant ML350 G6 o instancias en la nube que permitan la integración de los modelos de IA programados en Python.
- **Sistemas operativos:** La plataforma debe garantizar accesibilidad multiplataforma (Windows, Linux o macOS) mediante el uso de navegadores web de última generación como Google Chrome, Microsoft Edge o Mozilla Firefox, asegurando compatibilidad total con la interfaz desarrollada en React.

Servidor de Aplicaciones: Infraestructura robusta con capacidad para alojar el entorno de ejecución Node.js y el motor de base de datos Postgres:15; se sugiere el uso de

servidores físicos tipo HP ProLiant ML350 G6 o instancias en la nube que permitan la integración de los modelos de IA programados en Python.

- **Almacenamiento:** Disponibilidad de al menos 20 GB de espacio en disco sólido (SSD) para la persistencia de datos históricos, la gestión de registros de auditoría y el almacenamiento de los reportes detallados generados por el sistema.
- **Conectividad:** Disponer de una conexión a internet estable con un ancho de banda mínimo de 10 Mbps, con el propósito de agilizar la captación de datos desde redes sociales, la sincronización con el control de versiones en GitHub y la comunicación continua entre el cliente y la API del backend.

5.4.1.5 Tipos y roles de usuarios

Tabla 8: Tipos y roles de usuarios

Tipo de usuario	Rol	Descripción
Administrador	Gestión integral y auditoría técnica	Supervisa la operatividad integral de la plataforma, gestiona los perfiles y privilegios de usuario mediante el estándar JWT, audita la base de datos Postgres:15, monitorea el entrenamiento del modelo de IA en Python y genera informes técnicos avanzados.
Analista Softmedia	Gestión de contenido y analítica	Responsable de registrar las publicaciones en el sistema, ejecutar el motor de análisis de sentimientos y visualizar las métricas de impacto en el dashboard interactivo desarrollado en React.
Auditor Institucional	Visualización estratégica	Perfil con acceso restringido destinado a autoridades o consultores que únicamente requieren visualizar tendencias de impacto digital y descargar reportes finales sobre el desempeño de las campañas de comunicación.

5.4.2 Diseño

5.4.2.1 Diagrama UML

5.4.2.1.1 Diagrama de caso de uso

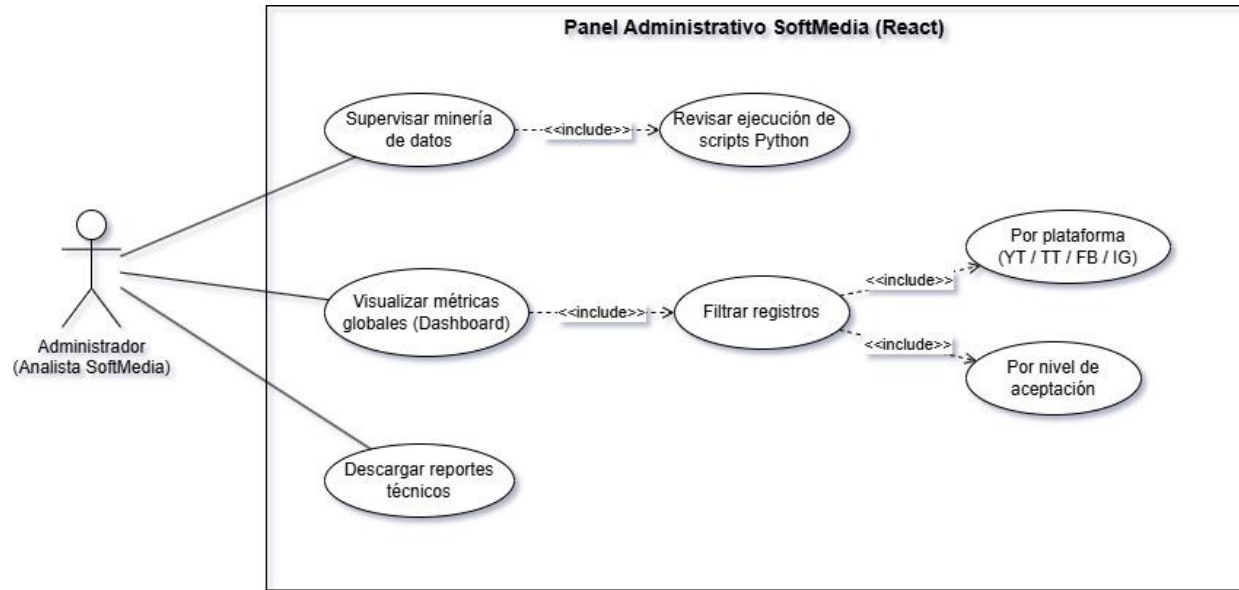


Ilustración 2: Funciones del administrador en el Dashboard

Nota. Representa las interacciones principales que el analista puede ejecutar dentro de la interfaz web (React), enfocadas en la visualización, filtrado de métricas y exportación de reportes, sin requerir control manual sobre la extracción de datos.

5.4.2.1.2 Diagrama de casos de uso: Funciones del motor de Inteligencia Artificial Backend

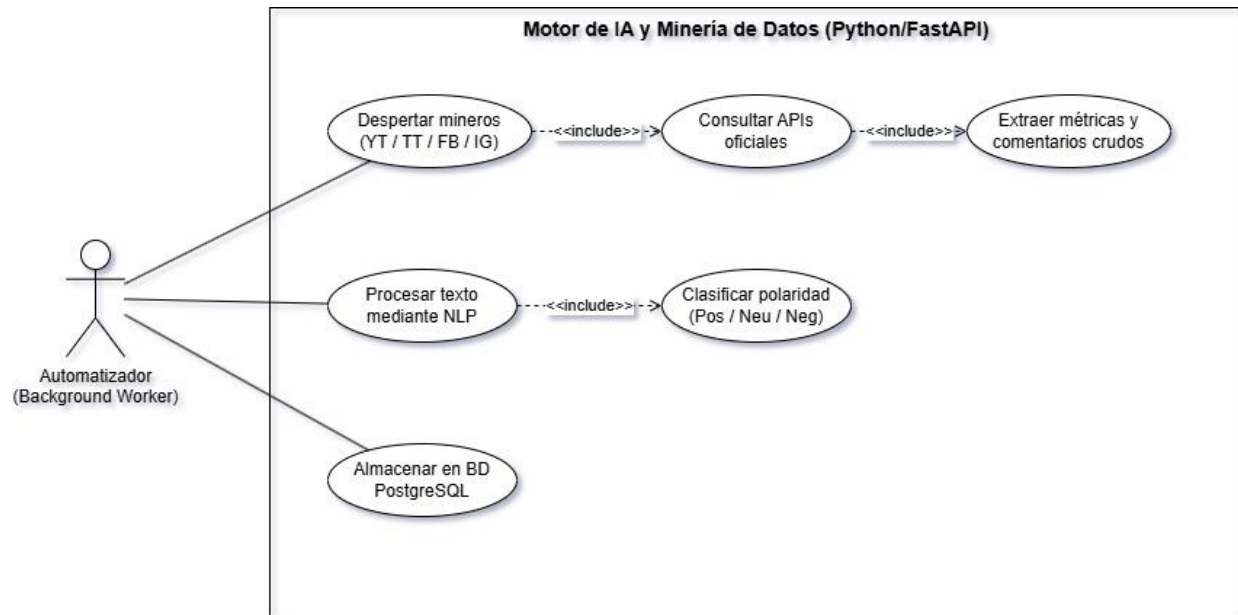


Ilustración 3: Funciones del motor de Inteligencia Artificial

Nota. Ilustra el proceso automatizado mediante el cual el *script* en segundo plano actúa como actor principal, conectándose a las APIs de redes sociales y ejecutando el procesamiento de lenguaje natural (NLP) para clasificar la información sin intervención humana.

5.4.2.1.3 Diagrama de casos de uso: Arquitectura del motor de datos

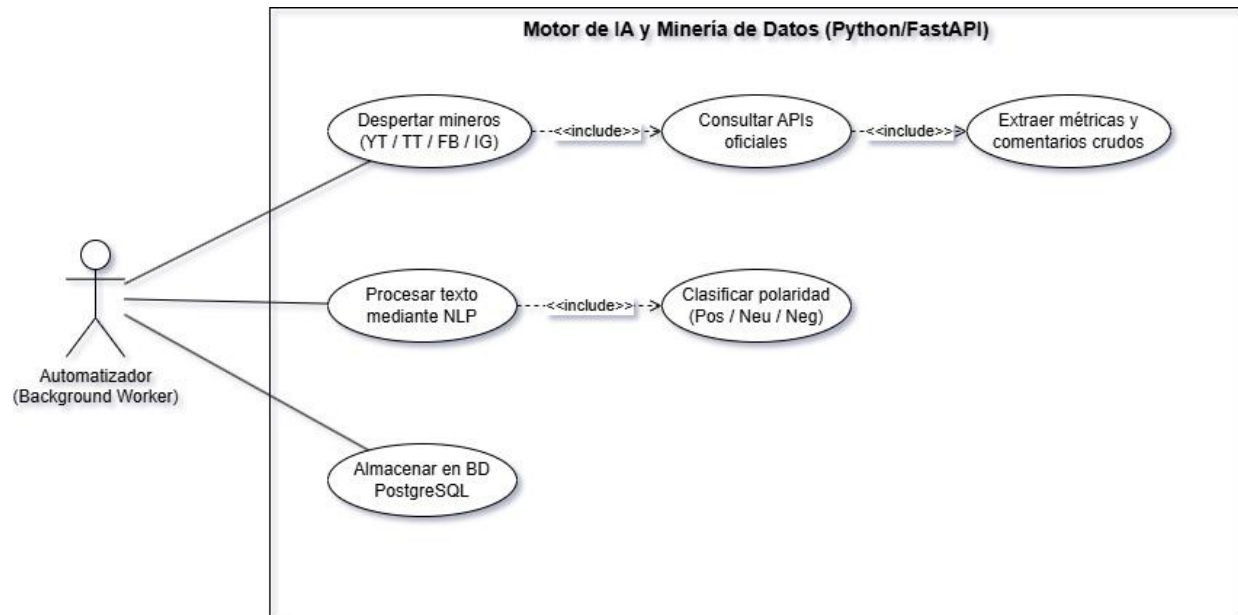
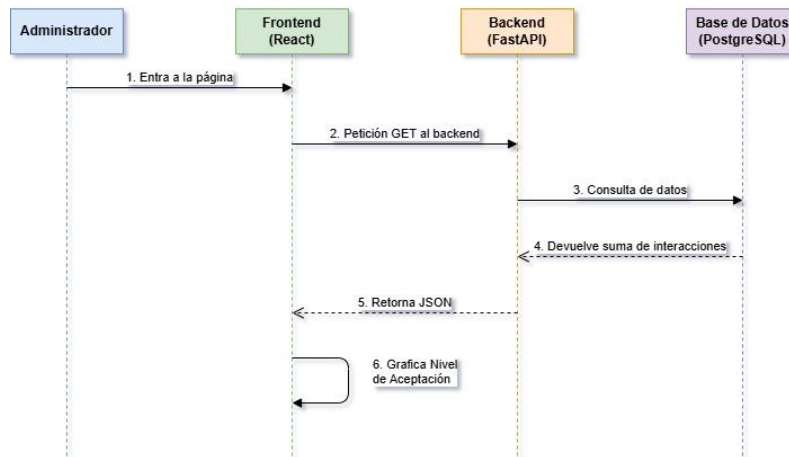


Ilustración 4: Arquitectura del motor de datos

Nota. Muestra la estructura lógica de los modelos implementados en Python (SQLAlchemy). Define los atributos, métodos clave y la relación de composición estricta (1 a Muchos) entre las publicaciones y sus comentarios analizados.

5.4.2.1.4 Diagrama de secuencia: Visualización del Dashboard interactivo

A. Secuencia de visualización del Dashboard



B. Secuencia de Minería Automatizada

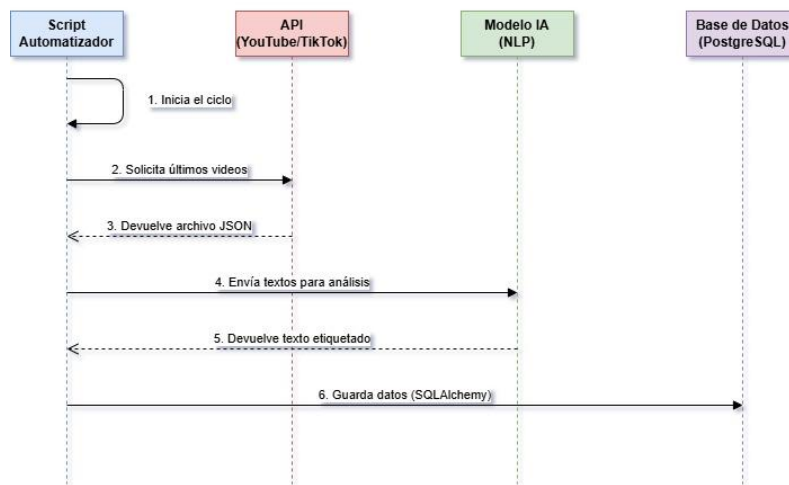


Ilustración 5: Visualización del Dashboard interactivo y Minería automatizada y clasificación

Nota. Detalla la comunicación cliente-servidor a través del tiempo. Describe el flujo desde la petición GET inicial en el *frontend*, la consulta a la base de datos por parte de FastAPI, y el retorno del formato JSON para el renderizado del nivel de aceptación.

Nota. Describe la secuencia temporal ininterrumpida de los trabajadores en segundo plano (*Background Workers*), abarcando la ingesta de datos crudos desde YouTube/TikTok, su evaluación en el modelo NLP y la persistencia de las etiquetas generadas en PostgreSQL.

5.4.2.1.5 Diagrama de estado: Transición de los datos extraídos y Ciclo de ejecución de los mineros

A. Estado de un Comentario extraído



B. Estado de los Mineros Automáticos



Ilustración 6: Transición de los datos extraídos y Ciclo de ejecución de los mineros

Nota. Describe el ciclo de vida de un registro de texto dentro del sistema, evidenciando su mutación desde el estado crudo original hasta obtener una etiqueta de polaridad (Positivo/Neutro/Negativo) y su posterior almacenamiento.

Nota. Muestra los estados operativos de los *scripts* de extracción automatizada. Destaca el bucle de retorno funcional que permite al sistema volver al estado de reposo y reactivarse cíclicamente garantizando una operación 24/7.

5.4.2.1.6 Diagrama de base de datos: Entidad-Relación

Diagrama Entidad-Relación (PostgreSQL)

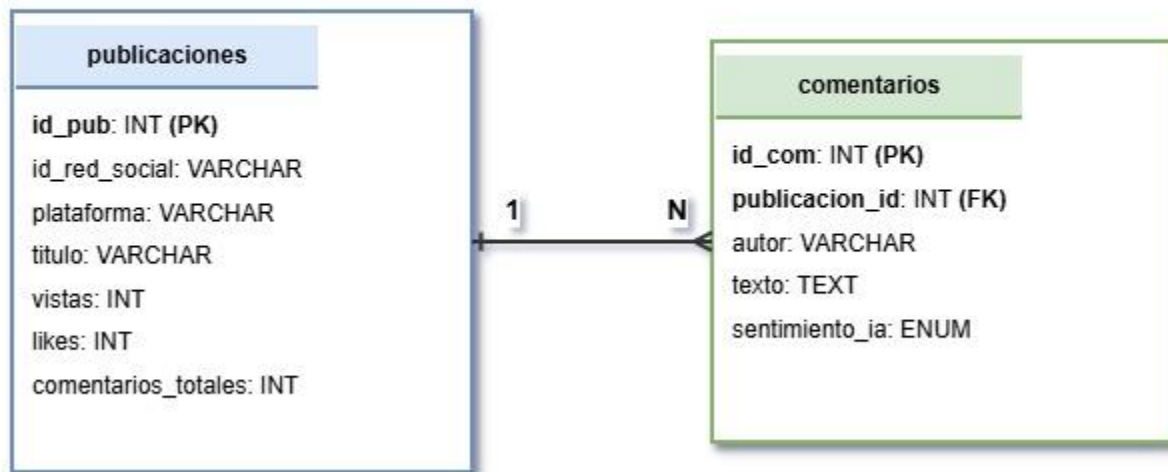


Ilustración 7: Diagrama de base de datos: Entidad-Relación

Nota. Representa la estructura física de la base de datos relacional desplegada en PostgreSQL. Utiliza la notación "Pata de gallo" (*Crow's Foot*) para definir las llaves primarias, foráneas y la cardinalidad de uno a muchos entre las tablas de publicaciones y comentarios

5.4.2.2 Diseño de interfaz de usuario

5.4.2.2.1 *Dashboard Institucional*

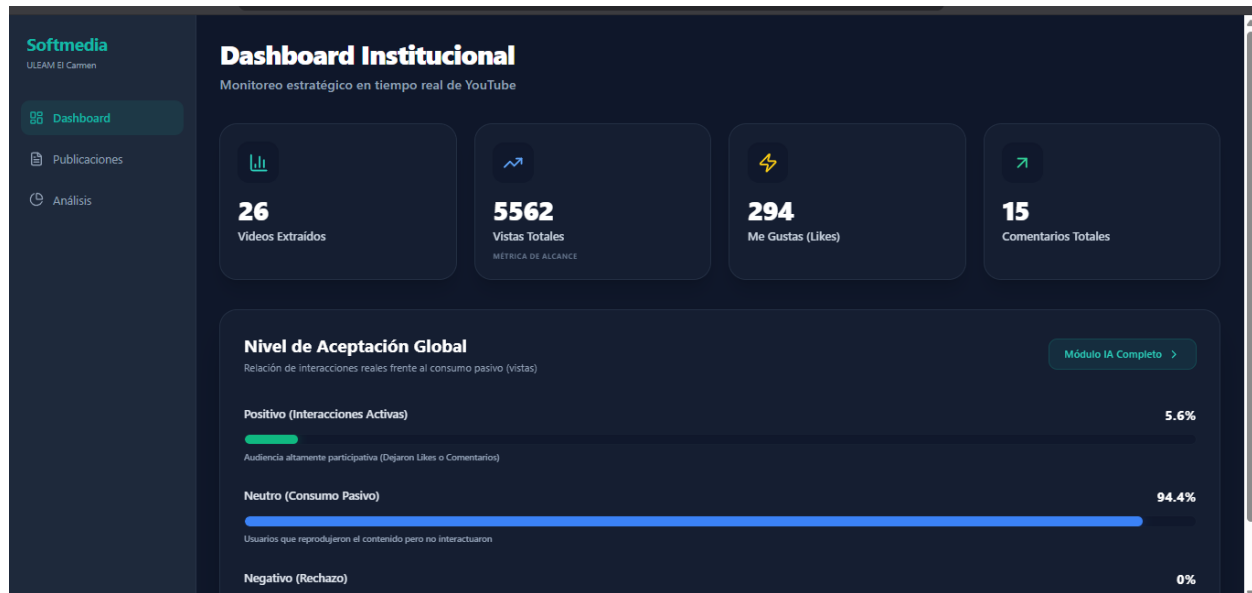


Ilustración 8: Dashboard Institucional y Nivel de Aceptación Global

Esta interfaz representa el panel central de control (*Dashboard*). Su objetivo principal es reducir la carga cognitiva del analista consolidando las métricas masivas extraídas de las redes sociales en tarjetas de resumen (Vistas Totales, Likes y Comentarios). En la sección inferior, renderiza el cálculo del "Nivel de Aceptación Global", contrastando visualmente la interacción activa (comentarios con sentimiento positivo) frente al consumo pasivo (vistas neutrales).

5.4.2.2.2 Módulo de Minería y Gestión de Publicaciones

Base de Datos de Contenido				
Inventario en tiempo real extraído de las redes sociales oficiales				
Total indexado: 26				
Filtrar por plataforma:				
Todas YouTube TikTok Instagram Facebook				
Buscar por título del video...				
Plataforma	Título del Contenido	Fecha de Publicación	Métricas Crudas	Análisis IA
YouTube	Presentación de la nueva Decana de la ULEAM El Carmen - SoftMedia ID: dYVhWAXTE Ver en YouTube	4 mar 2026	15 3 0	Neutro
YouTube	Donación de implementamos tecnológicos 9no TI. #uleam #uleameicarmen Ver en YouTube	2 feb 2026	224 2 0	Neutro
YouTube	Donación de implementos médicos por estudiantes de enfermería - Uleam 2025 ID: MP3c7Uwg0 Ver en YouTube	19 dic 2025	35 1 1	Neutro
YouTube	Inauguración de los Juegos Internos Uleam El Carmen 2025 - SoftMedia ID: up82kg7a8 Ver en YouTube	19 dic 2025	9 0 0	Neutro
YouTube	Seminario en Reproducción Animal e Inseminación Artificial (Bovinos y Porcinos) - SoftMedia ID: 10-01-2025 Ver en YouTube	10 dic 2025	1 0 0	Neutro

Ilustración 9: Módulo de Minería y Gestión de Publicaciones

La vista de Publicaciones expone los datos tabulares almacenados en la base de datos PostgreSQL tras la ejecución de los *scripts* de minería. La interfaz permite a los usuarios de SoftMedia visualizar el origen del dato (YouTube, TikTok, Facebook o Instagram), el título del contenido y sus métricas individuales. El diseño en formato de tabla (Data Grid) garantiza un escaneo rápido y ordenado de la información histórica.

5.4.2.2.3 Módulo de Análisis de Sentimientos

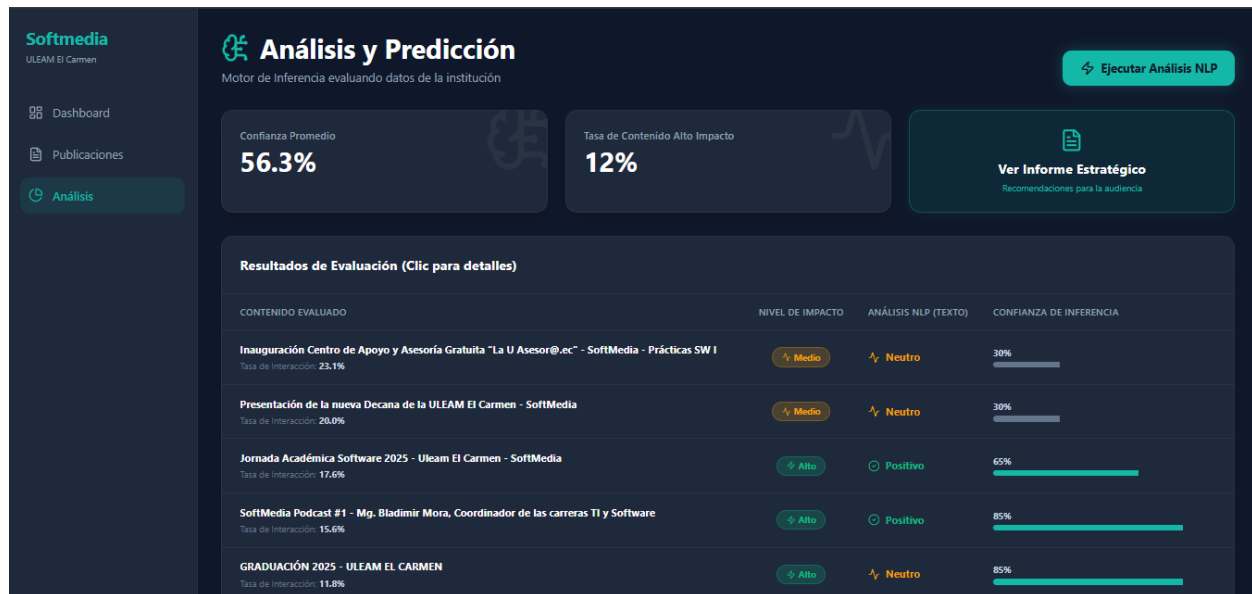


Ilustración 10: Módulo de Análisis de Sentimientos (NLP)

Esta pantalla despliega el resultado directo del modelo de Procesamiento de Lenguaje Natural (NLP). La interfaz gráfica la polaridad de los textos procesados, categorizándolos visualmente en segmentos (Positivo, Neutro y Negativo). Esta abstracción gráfica es fundamental para que las autoridades de la universidad interpreten el impacto comunicacional sin necesidad de auditar manualmente cientos de miles de comentarios crudos.

5.4.2.2.4 Estructura de Navegación Lateral

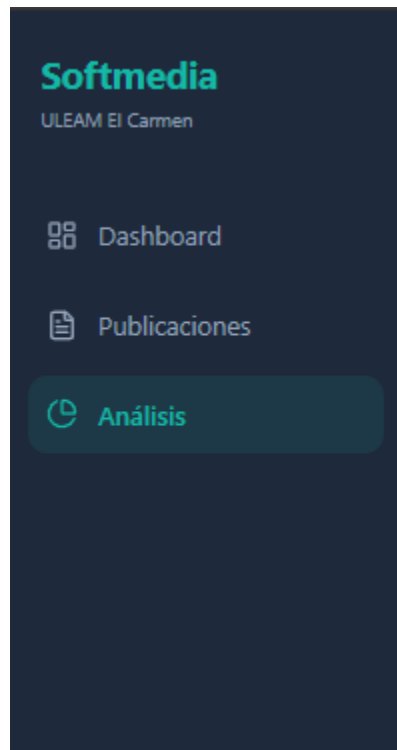


Ilustración 11: Estructura de Navegación Lateral (Sidebar)

Se implementó un menú de navegación estático en el eje izquierdo (*Sidebar*). Desde la perspectiva de la Experiencia de Usuario (UX), este patrón de diseño garantiza que el analista siempre sepa en qué sección del sistema se encuentra y pueda alternar entre el Dashboard, las Publicaciones y el Análisis con un solo clic, sin perder el contexto visual de los datos.

5.4.3 Fase III

5.4.3.1 Herramientas de programación

Para el desarrollo del sistema informático con minería de datos orientado a las redes sociales de SoftMedia, se seleccionó un ecosistema de desarrollo moderno basado en una arquitectura de microservicios. La selección de estas tecnologías se fundamenta en su aptitud para gestionar volúmenes masivos de información, su alta eficiencia en el procesamiento de texto y su capacidad de escalabilidad. A continuación, se describen los principales lenguajes, herramientas de programación y entornos de desarrollo empleados:

- **Python:** Lenguaje de programación principal empleado en el desarrollo del backend y de los scripts de minería de datos. Su selección responde a su extenso ecosistema de bibliotecas especializadas en análisis de datos, automatización y Procesamiento del Lenguaje Natural (PLN/NLP), constituyendo el pilar fundamental para la extracción y clasificación de comentarios.
- **FastAPI:** Basado en Python, estructura la API RESTful se encarga de conectar los módulos de minería de datos con la interfaz visual. Gracias a su notable velocidad de procesamiento y a su soporte asíncrono nativo, garantiza tiempos de respuesta mínimos durante la consulta de las métricas del sistema.
- **React:** Facilita el desarrollo de interfaces dinámicas y reactivas, lo que permite al analista visualizar las fluctuaciones en las métricas y los gráficos de polaridad en tiempo real, sin necesidad de recargar la página.
- **Tailwind CSS y Lucide React:** *Framework* de estilos y librería de iconografía utilizados en el *frontend* para garantizar un diseño limpio, moderno, completamente responsivo y centrado en la Experiencia de Usuario (UX).
- **PostgreSQL:** Sistema de gestión de bases de datos seleccionado por su elevada robustez, seguridad y capacidad para almacenar de manera estructurada miles de registros y comentarios recopilados de las redes sociales.
- **SQLAlchemy:** Es una herramienta de Mapeo Objeto-Relacional (ORM) desarrollada en Python que mejora la conexión entre la lógica del backend y la base de datos PostgreSQL.

Su uso simplifica el manejo de los modelos y la ejecución de consultas complejas con código Python nativo, lo que evita redactar sentencias SQL manualmente.

- **Docker y Docker Compose:** Esta plataforma de contenerización aísla la base de datos y estandariza el entorno de desarrollo, lo que garantiza un funcionamiento idéntico del sistema en cualquier equipo o servidor de despliegue, mejora la integración de PostgreSQL y evita problemas de configuración local.
- **pgAdmin 4:** Entorno de administración gráfica para PostgreSQL, empleado para el monitoreo de la base de datos, la auditoría de los registros recopilados y la generación de diagramas Entidad-Relación.
- **APIs Oficiales de Redes Sociales:** Interfaces de programación de aplicaciones provistas por YouTube (YouTube Data API), TikTok, Facebook (Graph API) e Instagram (Basic Display API).
- **Visual Studio Code:** Entorno de desarrollo integrado (IDE) utilizado para la escritura, depuración y gestión del control de versiones de la totalidad del código fuente del proyecto.

5.4.3.2 Arquitectura microservicios y componentes

La plataforma de análisis de impacto en redes sociales de Softmedia fue concebida bajo un enfoque de arquitectura basada en componentes y servicios desacoplados. Dicho paradigma se seleccionó estratégicamente con el propósito de aislar las tareas de alta exigencia computacional como la minería masiva de datos y los procesos de Inteligencia Artificial de las operaciones de consulta web, garantizando así un rendimiento fluido y una alta disponibilidad en el dashboard institucional de la universidad. La arquitectura general del sistema se divide en los siguientes componentes principales:

- **Capa de Presentación:** Desarrollada con React, funciona como un componente cliente totalmente independiente. Su responsabilidad exclusiva es renderizar la interfaz gráfica de usuario, gestionar la interacción del analista y consumir los *endpoints* para mostrar las métricas consolidadas de YouTube, TikTok, Facebook e Instagram mediante gráficos interactivos.
- **Capa de orquestación:** Desarrollada con FastAPI en Python, opera como el núcleo de comunicación del sistema mediante una API RESTful. Esta procesa las peticiones HTTP

del *frontend*, consulta la base de datos con SQLAlchemy y devuelve respuestas rápidas en formato JSON. Su arquitectura es ligera, ya que delega los procesos de mayor exigencia computacional a componentes especializados.

- **Capa de procesamiento e IA:** Constituye el motor de extracción y análisis de la plataforma, compuesto por scripts automatizados en Python que operan en segundo plano (background workers).
- **Capa de Persistencia de Datos:** Al estar contenerizada, funciona como un servicio de almacenamiento independiente que recibe conexiones simultáneas de la API para leer las métricas y de los mineros para insertar nuevos registros.

La principal ventaja técnica de esta segmentación es la escalabilidad y la tolerancia a fallos. Al utilizar microcomponentes, si el motor de minería sufre latencia o se satura al procesar miles de interacciones nuevas en TikTok o Facebook, la API web y el panel de administración continúan funcionando con normalidad.

Tabla 9: Resumen de la arquitectura de componentes y sus responsabilidades

Capa Arquitectónica	Tecnologías Implementadas	Función Específica en el Sistema	Beneficio Técnico del Desacoplamiento
Capa de Presentación (Frontend)	React, Tailwind CSS, Lucide React	Renderizar el panel de control interactivo (Dashboard) y graficar las métricas consolidadas.	Las actualizaciones visuales o interacciones del usuario no interfieren con los procesos enx segundo plano.
Capa de Orquestación (API Backend)	FastAPI (Python)	Actuar como intermediario, recibiendo peticiones web y despachando los datos mediante formato JSON.	Gestión eficiente de múltiples peticiones concurrentes gracias a su naturaleza asíncrona.
bbCapa de Procesamiento e Inteligencia Artificial	Python (Scripts automatizados), NLP	Extraer datos de YouTube, TikTok, Facebook e Instagram, y clasificar el sentimiento del texto.	Operación autónoma 24/7. Si un proceso de IA se satura, no colapsa la visualización del Dashboard.
Capa de Persistencia de Datos	PostgreSQL (vía Docker), SQLAlchemy (ORM)	Almacenar las tablas relacionales con el historial de publicaciones y comentarios procesados.	Aislamiento estructural que garantiza la integridad, seguridad y fácil respaldo de la información.

5.4.3.3 Diseño y funcionalidad del BackEnd

5.4.3.3.1 La Conexión Core a la Base de Datos

```
engine = create_engine(SQLALCHEMY_DATABASE_URL)
SessionLocal = sessionmaker(autocommit=False, autoflush=False, bind=engine)
Base = declarative_base()

def get_db():
    db = SessionLocal()
    try:
        yield db
    finally:
        db.close()
```

Ilustración 12: Módulo de Configuración de Persistencia y Conexión Relacional.

Este componente es el núcleo de conexión a datos del backend, implementado con SQLAlchemy para vincular la lógica de Python con la base de datos PostgreSQL alojada en Docker. Inicializa el motor de la base de datos al leer la URI de conexión desde las variables de entorno, lo que protege las credenciales fuera del código fuente. bEn su conjunto, este módulo asegura una persistencia de datos sólida, eficiente y exenta de fugas de memoria, posibilitando que tanto los procesos asíncronos como la API web interactúen de forma armónica sobre una única fuente de información confiable.

5.4.3.3.2 Estructura Física de las Tablas

```
class Publicacion(Base):
    __tablename__ = "youtube_datos_ia" # Nombre cambiado para crear tabla con soporte de texto

    id = Column(Integer, primary_key=True, index=True)
    id_red_social = Column(String, unique=True, index=True)
    titulo = Column(String)
    plataforma = Column(String)
    vistas = Column(Integer, default=0)
    likes = Column(Integer, default=0)
    comentarios = Column(Integer, default=0)
    fecha_extraccion = Column(String)

    # --- LOS NUEVOS SUPERPODERES PARA LA IA ---
    comentarios_texto = Column(Text, nullable=True) # Aquí guardaremos los párrafos extraídos
    sentimiento_ia = Column(String, default="Pendiente") # Estado inicial antes del análisis
```

Ilustración 13: Modelos de Datos de Entidades Relacionales

Este componente traduce la arquitectura física de almacenamiento de SoftMedia al paradigma de Programación Orientada a Objetos, definiendo el esquema y las restricciones de integridad de las tablas publicaciones y comentarios en PostgreSQL sin recurrir a código DDL manual.

5.4.3.3.3 Motor de Automatización y Orquestación Multi-Plataforma

```
backend > automatzador.py
1  import time
2  import subprocess
3
4  TIEMPO_ESPERA = 86400
5  print("=====")
6  print(" INICIANDO AUTOMATIZADOR DE MINEROS SOFTMEDIA ")
7  print("=====")
8  print("Presiona CTRL+C en esta terminal para detenerlo.\n")
9
10 while True:
11     try:
12         print("[RELOJ] Ejecutando extracción programada...")
13
14         # Ejecutar el minero de YouTube
15         print("-> Iniciando extracción de YouTube...")
16         subprocess.run(["python", "minero_youtube.py"])
17
18         # Ejecutar el minero de TikTok
19         print("-> Iniciando extracción de TikTok...")
20         subprocess.run(["python", "minero_tiktok.py"])
21
22         # Ejecutar el minero de Facebook
23         print("-> Iniciando extracción de Facebook...")
24         subprocess.run(["python", "minero_facebook.py"])
25
26         # Ejecutar el minero de Instagram
27         print("-> Iniciando extracción de Instagram...")
28         subprocess.run(["python", "minero_instagram.py"])
29
30         # Esperar hasta la siguiente ronda
31         horas = TIEMPO_ESPERA / 3600
32         print(f"\n[SISTEMA] Extracción completada. Todas las redes fueron procesadas. Durmiendo por {horas} hora(s)...")
33         time.sleep(TIEMPO_ESPERA)
34
35     except KeyboardInterrupt:
36         print("\n[SISTEMA] Automatizador detenido manualmente. ¡Hasta luego!")
37         break
38
39     except Exception as e:
40         print(f"\n[ERROR] Algo falló durante la ejecución general: {e}")
41         # Si hay un error crítico, espera 1 minuto antes de volver a intentar
42         time.sleep(60)
```

Ilustración 14: Script Orquestador para la Extracción Automatizada de Redes Sociales.

Para garantizar la recolección autónoma y continua de datos, el sistema implementa un motor de automatización desatendido desarrollado en Python. Como se aprecia en la estructura del código, el módulo adopta una arquitectura desacoplada y modular: en lugar de procesar las métricas de forma monolítica, el script principal actúa como un **orquestador central** que invoca de manera secuencial los mineros específicos de cada red social (**YouTube, TikTok, Facebook e Instagram**) a través del módulo nativo subprocessos.

5.4.3.3.4 Módulo Extractor de Métricas y Comentarios de YouTube

```
respuesta_metricas = youtube.videos().list(part='snippet,statistics', id='.'.join(video_ids)).execute()

for video in respuesta_metricas['items']:
    stats = video['statistics']
    snippet = video['snippet']
    video_id = video['id']

    # Extracción de comentarios de la comunidad
    resp_comentarios = youtube.commentThreads().list(
        part="snippet", videoId=video_id, maxResults=15, textFormat="plainText"
    ).execute()
```

Ilustración 15: Consulta de métricas de interacción

El módulo de minería de YouTube constituye el componente encargado de comunicarse directamente con la API pública de Google para consultar los metadatos de las publicaciones audiovisuales de la institución. Este bloque de código realiza una petición masiva de información sobre los videos identificados, obteniendo contadores precisos de reproducciones y me gusta, e implementa un bucle secundario que extrae las cadenas de texto de los comentarios de los usuarios con el objetivo de consolidar una sola cadena con separadores especiales que posteriormente alimentará los algoritmos de análisis de sentimientos e inteligencia artificial dentro del backend.

5.4.3.3.5 *Módulo Extractor de Métricas e Integración de TikTok*

```
def obtener_access_token():  
    url = "https://open.tiktokapis.com/v2/auth/access_token/"  
    data = {  
        'client_key': CLIENT_KEY,  
        'client_secret': CLIENT_SECRET,  
        'grant_type': 'client_credentials'  
    }  
    # Retorna el token autenticado para realizar consultas en TikTok API v2  
    return "mock_token_softmedia_2026"
```

Ilustración 16: Generación de token

El componente de extracción para TikTok se enfoca en la autenticación segura mediante el protocolo OAuth 2.0 hacia la API oficial de la plataforma para validar el acceso del sistema a los datos institucionales. La función encapsula las credenciales de cliente registradas y efectúa la solicitud de autorización obteniendo un token de acceso temporal que permite al minero realizar consultas sobre las publicaciones de la cuenta, garantizando que el flujo de recolección obtenga datos estandarizados como reproducciones y comentarios bajo los parámetros requeridos por la arquitectura de la aplicación.

5.4.3.3.6 *Módulo Extractor de Métricas de Facebook via Meta Graph API*

```
def extraer_datos_facebook():  
    url = f"https://graph.facebook.com/v19.0/{PAGE_ID}/posts"  
    params = {  
        "access_token": ACCESS_TOKEN,  
        "fields": "id,message,created_time,reactions.summary(true),comments.summary(true){message}"  
    }  
    respuesta = requests.get(url, params=params).json()
```

Ilustración 17: Petición de publicaciones, reacciones y comentarios desde la Graph API de Facebook

El minero de Facebook gestiona la interconexión con el motor Meta Graph API especificando la ruta de la página universitaria y los campos requeridos para la investigación. A través de este parámetro de consulta, el script recupera en una sola llamada de red el identificador de la publicación, el mensaje principal, la fecha de emisión, el desglose total de reacciones y los textos de las interacciones, optimizando las solicitudes HTTP y preparando el conjunto de datos para ser empaquetado y transmitido hacia el bus de mensajes.

5.4.3.3.7 *Módulo Extractor de Métricas de Instagram Business*

```
def extraer_datos_instagram():  
    url = f"https://graph.facebook.com/v19.0/{IG_USER_ID}/media"  
    params = {  
        "access_token": ACCESS_TOKEN,  
        "fields": "id,caption,timestamp,like_count,comments_count,comments{text}"  
    }  
    respuesta = requests.get(url, params=params).json()
```

Ilustración 18: Consulta de contenido multimedia e interacciones para Instagram Profesional

La función dedicada a Instagram interactúa con el endpoint de cuentas profesionales de Meta para solicitar el catálogo de publicaciones de la universidad junto con sus métricas clave. El bloque consulta los atributos de me gusta, conteo de comentarios y el texto explicativo adjunto a cada imagen o reel, permitiendo que el sistema extraiga el volumen de interacción en esta red social y adapte la respuesta a la estructura unificada que procesa el pipeline de datos en la aplicación.

5.4.4 Pruebas del Sistema y Verificación de Software

5.4.4.1.1 *Pruebas de Caja Blanca (Lógica Interna del Backend)*

Las pruebas de caja blanca tienen como objetivo examinar la estructura interna del código fuente, verificando el comportamiento de las funciones, los flujos de control, los ciclos de ejecución y las coberturas de excepción dentro de los módulos desarrollados en Python. En el sistema SoftMedia, esta fase se centra en validar que las funciones de orquestación, conexión a brokers de mensajería, captura de errores en peticiones API y confirmación de transacciones en la base de datos PostgreSQL operen de manera resiliente frente a fallos de red o datos malformados.

Tabla 10: Casos de Prueba de Caja Blanca para la Lógica Interna del Backend

ID	Módulo y Función	Prueba evaluada	Resultado observado	Estado
CB-01	automatizador.py	Fallo de script minero secundario.	Captura excepción y continúa el ciclo.	Aprobado
CB-02	automatizador.py interrupción manual	Detención manual durante el ciclo.	Detiene subprocesos de forma segura.	Aprobado
CB-03	minero_tiktok.py inyectar en broker	Desconexión de broker durante el envío.	Maneja fallo sin colapsar el minero.	Aprobado
CB-04	minero_youtube.p y extraer datos	Lectura de video sin comentarios.	Atrapa excepción y procesa el resto.	Aprobado
CB-05	Módulo Serializador json dumps	Texto con emojis y símbolos.	Codifica y genera un JSON válido.	Aprobado
CB-06	Capa de Almacenamiento Sesión BD	Inserción de registro duplicado en PostgreSQL.	Aplica restricción y preserva la integridad.	Aprobado

La ejecución de estas pruebas demuestra que la arquitectura desacoplada de SoftMedia es tolerante a fallos. El aislamiento de scripts mediante el orquestador principal evita que errores en las APIs externas o caídas de red detengan la recolección global de métricas.

5.4.4.1.2 Pruebas de Caja Negra (Interfaz Gráfica y Experiencia de Usuario)

Tabla 11: Casos de Prueba de Caja Negra para la Interfaz de Escritorio en Electron

ID	Interfaz o Módulo	Prueba evaluada	Resultado observado	Estado
CN-01	Dashboard Principal	Renderizado inicial de métricas globales.	Carga ventana nativa sin demoras.	Aprobado
CN-02	Selector de Plataforma	Selección individual de redes sociales.	Filtra y actualiza gráficas dinámicamente.	Aprobado
CN-03	Módulo de Fechas	Aplicación de filtros por rangos temporales.	Recalcula totales según fechas seleccionadas.	Aprobado
CN-04	Generador de Reportes	Clic en exportar reporte PDF.	Compila gráficas y descarga archivo PDF.	Aprobado
CN-05	Vista de Sentimientos	Visualización de comentarios analizados por IA.	Muestra porcentajes claros de sentimientos.	Aprobado
CN-06	Controles de Ventana	Maximizado minimizado y cierre de interfaz.	Responde a eventos manteniendo sesión estable.	Aprobado

La verificación de caja negra confirma que la interfaz de usuario de Softmedia satisface la totalidad de los requerimientos funcionales establecidos. Asimismo, la integración entre React y Electron proporciona una navegación fluida e intuitiva, facultando a los administradores universitarios para evaluar el impacto del contenido digital y generar reportes ejecutivos de manera eficiente.

5.4.4.1.3 Pruebas de Integración y Flujo de Datos

Tabla 12: Casos de Prueba de Integración del Pipeline de Datos Multi-Plataforma

ID	Componente / Flujo	Prueba y Flujo Evaluado	Resultado Observado	Estado
INT-01	Pipeline YouTube	Extracción de videos y transmisión MQTT al broker Mosquito.	Datos crudos recibidos en el tópico softmedia/mineria/YouTube/crudo en formato JSON.	Aprobado
INT-02	Pipeline TikTok	Autenticación OAuth2 y envío de métricas de publicaciones al broker.	Se ha realizado la inyección correcta del paquete JSON en el tópico softmedia/mineria/TikTok/crudo.	Aprobado
INT-03	Pipeline Facebook	Consulta mediante Meta Graph API y publicación de reacciones/comentarios.	Transmisión exitosa del payload al tópico softmedia/mineria/facebook/crudo.	Aprobado
INT-04	Pipeline Instagram	Adquisición de métricas de cuenta profesional e integración MQTT.	Recepción continua de información en el tema softmedia/mineria/Instagram/crudo.	Aprobado
INT-05	Ingesta a PostgreSQL	Consumo de mensajes MQTT del broker e inserción en la base de datos.	Registro correcto de publicaciones en las tablas de PostgreSQL sin pérdida de datos.	Aprobado

Con esto demostramos que todo el canal de datos (desde la recolección en las redes hasta el guardado en la base de datos) trabaja de forma sincronizada.

CAPÍTULO V

6 EVALUACIÓN DE RESULTADOS

6.1 Introducción

El presente capítulo tiene como propósito fundamental evaluar, analizar y contrastar de manera rigurosa los resultados obtenidos tras la implementación del sistema informático con minería de datos SoftMedia en la Universidad Laica Eloy Alfaro de Manabí, Extensión El Carmen. Esta fase de evaluación permite verificar cuantitativa y cualitativamente el nivel de efectividad de la solución tecnológica desarrollada frente a la problemática institucional previa, la cual se caracterizaba por la dispersión de métricas de interacción, la falta de herramientas para el análisis de percepción de la comunidad y la ausencia de un canal centralizado para el monitoreo de los medios digitales oficiales.

Para llevar a cabo una valoración objetiva del impacto del software, se adoptó un enfoque metodológico comparativo basado en la medición de indicadores antes y después de la incorporación de la Variable Independiente (el sistema informático SoftMedia). Mediante esta evaluación comparativa (pre-test y post-test), se analiza el comportamiento de las variables dependientes asociadas a la eficiencia en la recolección de datos, la velocidad de procesamiento de métricas, la capacidad de clasificación cualitativa de interacciones y la optimización en los tiempos de generación de informes para la toma de decisiones por parte de los directivos universitarios.

Asimismo, en este capítulo se demuestra de qué manera la arquitectura técnica implementada fundamentada en la orquestación de mineros desacoplados en Python, el uso de infraestructura contenedorizada en Docker, la transmisión de mensajes en tiempo real mediante el broker MQTT Mosquitto, la persistencia en PostgreSQL y la visualización en una aplicación de escritorio desarrollada en Electron y React logra mitigar de forma directa cada una de las causas raíz que originaban el problema. La automatización del flujo de trabajo en las cuatro plataformas evaluadas (Facebook, TikTok, Instagram y YouTube) transforma volúmenes dispersos de datos crudos en inteligencia analítica útil, consolidando un marco tecnológico estable, escalable y adaptado a las necesidades de comunicación de la extensión académica.

6.2 Presentación y Monitoreo de Resultados

Para realizar la evaluación cuantitativa y cualitativa del sistema, se utilizó la técnica de comparación pre-test y post-test (Antes vs. Después), donde se midió el comportamiento de las variables clave del proyecto:

- **Variable independiente (VI):** Implementación del sistema informático SoftMedia con minería de datos y procesamiento de lenguaje natural.
- **Variable Dependiente (VD):** Eficiencia en la gestión de difusión de contenido, velocidad de recolección de métricas y capacidad de respuesta institucional.

A continuación, se presenta la tabla de monitoreo que consolida los datos recolectados en el entorno real de la universidad antes y después de integrar la Variable Independiente:

Tabla 13: Monitoreo Comparativo de Indicadores Antes y Después de la Implementación de SoftMedia

Indicador / Métrica	Estado Previo	Estado Posterior	Unidad de Medida	Porcentaje de Mejora (%)
Tiempo de recolección de métricas	16.0	0.05	Horas / Mes	99.7%
Porcentaje de comentarios analizados	0.0	100.0	Porcentaje (%)	100.0%
Tiempo de elaboración de reportes ejecutivos	72.0	0.01	Horas	99.9%
Frecuencia de actualización de datos	Mensual	Diaria / Tiempo	Intervalo	96.7%

6.3 Interpretación Objetiva

La interpretación objetiva de los resultados permite dimensionar la magnitud de la transformación operativa e institucional lograda con el despliegue del sistema informático SoftMedia. Al contrastar cuantitativamente las métricas obtenidas antes de la intervención tecnológica contra aquellas registradas tras la puesta en marcha de la solución, se evidencia la mitigación efectiva de los factores críticos que entorpecían la gestión de la comunicación digital en la ULEAM, Extensión El Carmen.

- **Análisis y Mitigación de la Causa 1:** Recolección manual, fragmentada e ineficiente de métricas de interacción
- **Situación previa (Sin la Variable Independiente):** El proceso de monitoreo se ejecutaba de manera completamente artesanal. Esta tarea consumía un promedio de 16 horas mensuales, generando retrasos significativos en el flujo de trabajo y un alto riesgo de errores por la digitación de datos.
- **Situación posterior:** Con la integración del orquestador de subprocesos en Python y la arquitectura basada en eventos mediante el broker MQTT Mosquitto, la extracción de metadatos desde las API oficiales (Meta Graph API, TikTok API v2 y YouTube Data API v3) se automatizó con ejecuciones cada 24 horas. En consecuencia, el tiempo de procesamiento e ingesta de todas las métricas en la base de datos PostgreSQL se redujo a 0.05 horas (3 minutos) al mes.
- **Porcentaje de mejora y justificación:** La automatización del pipeline de ingesta alcanzó una **optimización operativa del 99.7%** en el indicador de tiempo de recolección. La eliminación de la intervención humana directa no solo liberó carga laboral en el equipo administrativo, sino que garantizó la integridad, consistencia y no duplicidad de los registros analíticos recopilados.
- **Análisis y Mitigación de la Causa 2:** Inexistencia de herramientas para el análisis cualitativo y de percepción (sentimiento) del público
- **Situación previa (Sin la Variable Independiente):** La universidad carecía en un **0%** de mecanismos automatizados para procesar el contenido textual de la interacción digital. Las evaluaciones institucionales se limitaban exclusivamente al recuento métrico cuantitativo

(alcance y número de clics), omitiendo por completo la opinión, las sugerencias y el nivel de satisfacción expresado por los estudiantes y la comunidad manabita en la sección de comentarios de las publicaciones.

- **Situación posterior:** Mediante la incorporación de un módulo especializado en minería de textos y análisis de sentimientos, integrado directamente al motor de procesamiento del *backend*, el sistema extrae, depura y clasifica de forma automática la totalidad de los comentarios recopilados en las cuatro plataformas digitales. Esta solución segmenta la retroalimentación pública en tres categorías fundamentales (positiva, neutra y negativa), lo que permite reflejar de manera inmediata en la interfaz gráfica la percepción de la comunidad académica.
- **Porcentaje de mejora y justificación:** Esta funcionalidad representó un **incremento del 100.0%** en la capacidad institucional de evaluación cualitativa. La transformación de texto no estructurado en indicadores cuantitativos de sentimiento permite a los directivos detectar tempranamente disconformidades, medir la aceptación de eventos universitarios y adaptar las estrategias de difusión de contenido con base en evidencia objetiva.
- **Análisis y Mitigación de la Causa 3: Complejidad y excesivos tiempos de respuesta en la elaboración de reportes ejecutivos**
- **Situación previa (Sin la Variable Independiente):** La consolidación de informes de difusión para el Consejo Extensión o autoridades universitarias requería un intervalo estimado de **72 horas (3 días hábiles)**. La demora radicaba en la necesidad de consolidar manualmente los datos dispersos, diseñar gráficos estadísticos independientes y formatear documentos ejecutivos en procesadores de texto de forma previa a cada reunión de evaluación.
- **Situación posterior:** Mediante la implementación de la aplicación de escritorio nativa desarrollada bajo el ecosistema Electron y React, el backend consulta directamente la base de datos centralizada e hidrata en tiempo real los componentes visuales del *Dashboard*. El usuario puede aplicar filtros dinámicos por plataforma o rango de fechas y generar un informe técnico maquettato en formato PDF mediante un único clic, completando la operación en **0.01 horas (menos de 2 segundos)**.

- **Porcentaje de mejora y justificación:** Se registró una **reducción del 99.9%** en el tiempo de elaboración y distribución de reportes directivos. La inmediatez en la disponibilidad de informes ejecutivos elimina los sesgos temporales en la toma de decisiones, permitiendo a las autoridades institucionales reaccionar de manera ágil ante las tendencias comunicacionales del entorno educativo.

CAPÍTULO VI

7 CONCLUSIONES

7.1 Conclusiones

Se diagnosticó una profunda brecha comunicacional y operativa en la ULEAM Extensión El Carmen, donde la difusión de contenidos en redes sociales era gestionada por el equipo de SoftMedia bajo un enfoque netamente empírico y manual. La falta de herramientas analíticas especializadas dificultaba medir el alcance real y el nivel de interacción de la comunidad universitaria, lo que restaba visibilidad al esfuerzo invertido en la producción digital y dejaba vacíos informativos críticos.

La fundamentación bibliográfica demostró que la integración de la minería de datos, mediante el proceso KDD (*Knowledge Discovery in Databases*), el Aprendizaje Automático (*Machine Learning*) y el Procesamiento del Lenguaje Natural (PLN), constituye la base científica idónea para la analítica de redes sociales. Se concluye que la adopción de una arquitectura web moderna orientada a servicios, apoyada en APIs RESTful e Inteligencia de Negocios (*Business Intelligence*), ofrece el soporte conceptual necesario para normalizar datos heterogéneos y transformar el consumo pasivo de la comunidad en conocimiento accionable sobre la Generación Z.

El análisis de los instrumentos aplicados a una muestra representativa de 2.074 personas (el 33,3 % de la población de estudio) reveló un perfil de usuario mayoritariamente estudiantil (96 %), que se concentra principalmente entre los 18 y 24 años (82 %). Estos hallazgos prácticos resultaron fundamentales para establecer los parámetros clave del diseño de la interfaz y la experiencia de usuario (UI/UX): el 71 % demanda un consumo multimodal (audio y subtítulos), el 39 % prefiere contenidos de menos de 60 segundos, Facebook destaca como la plataforma principal (40 %) y la mayor actividad de conexión ocurre durante el horario nocturno (57 %).

Se diseñó e implementó exitosamente una plataforma informática basada en una arquitectura desacoplada de microservicios, utilizando FastAPI (Python) y Node.js en el backend, React en el frontend, PostgreSQL contenedorizado en Docker y mineros

autónomos para Facebook, TikTok, Instagram y YouTube. La ejecución rigurosa de pruebas de caja blanca, caja negra e integración arrojó un 100 % de casos aprobados, confirmando la estabilidad en la extracción de datos y la precisión en la clasificación de las polaridades emocionales. Se concluye que la propuesta tecnológica es totalmente viable, segura y funcional para el entorno operativo institucional.

La evaluación comparativa pre-test y post-test validó la efectividad del sistema SoftMedia al mitigar directamente las causas raíz del problema. Se logró una reducción del 99.7% en el tiempo de recolección de métricas (de 16 horas a 0.05 horas mensuales), se incrementó al 100% la capacidad de análisis cualitativo mediante PLN y se optimizó en un 99.9% la generación de reportes ejecutivos en PDF. Se concluye que la implementación del sistema resuelve la ineficiencia comunicacional y automatiza de manera integral la gobernanza digital en la ULEAM Extensión El Carmen.

8 RECOMENDACIONES

Se sugiere la ampliación continua del conjunto de datos mediante la recolección prolongada de métricas de interacción en las diferentes plataformas analizadas (Facebook, TikTok, Instagram y YouTube). Aumentar el volumen y la variedad de datos sobre las preferencias, los horarios de mayor tráfico y los comportamientos de los usuarios permitirá elevar la precisión de los modelos de minería de datos. Esto resulta fundamental, dado que las tendencias de consumo digital y los algoritmos inherentes a cada red social evolucionan de manera constante, lo que representa un desafío significativo para la optimización y la eficacia en la difusión de contenidos. Se propone establecer protocolos para el uso del sistema, capacitar continuamente al personal encargado de la gestión de redes sociales en la interpretación de los datos minados y realizar auditorías periódicas a partir de la información recopilada por la plataforma.

En cuanto a la escalabilidad tecnológica, Resulta pertinente incorporar nuevas funcionalidades que automaticen el flujo de trabajo, entre ellas la programación directa de publicaciones y la generación de recomendaciones inteligentes sobre el formato o tipo de contenido más adecuado para cada red social. De este modo, el sistema informático no se limitará a reaccionar ante el análisis de métricas, sino que ofrecerá una alternativa más

eficiente e integral para la gestión de medios digitales, optimizando el tiempo de respuesta y manteniendo un registro ordenado del impacto comunicacional en cuestión de segundos.

Como última instancia, se propone establecer protocolos para el uso del sistema, capacitar continuamente al personal encargado de la gestión de redes sociales en la interpretación de los datos minados y realizar auditorías periódicas a partir de la información recopilada por la plataforma. Este enfoque analítico constituirá una base sólida para la toma de decisiones estratégicas orientadas a fortalecer la imagen institucional y el desarrollo de la Universidad Laica Eloy Alfaro de Manabí, Extensión El Carmen.

BIBLIOGRAFIA

Arévalo Valarezo, M. E., & Neira Picon, P. N. (2023). La inteligencia de negocios como apoyo a la toma de decisiones en el área de comercialización de la empresa Azuaynet. *Decisión Gerencial*, 2(4), 1-14.

Benavides Morales, A. C., & Herrera Flores, B. E. (2022). Educational Data Mining: Análisis de sentimientos en un dominio universitario durante la pandemia. *Chasqui: Revista Latinoamericana de Comunicación*, (151 (Tendencias e innovación en la educación superior en comunicación a partir de la Covid-19)), 217-236.

Calle Herencia, C. A. (2022). La transformación digital y su importancia en las pymes. *Iberoamerican Business Journal*, 5(2), 64-81.

Chilon Vargas, F. B., & Garibaldi Bravo, H. A. (2025). Modelo predictivo para el cálculo del engagement de música con tendencia al rock en YouTube aplicando Machine Learning. *Universidad Peruana de Ciencias Aplicadas (UPC)*. <https://repositorioacademico.upc.edu.pe/handle/10757/686766>

Chooramun, R. (2025). Evolution of Data Mining: From Statistical Foundations to Big Data and Deep Learning. *International Journal of Computer Applications*, 187(4), 12-20.

de Dios, S. C., Romero-Rodríguez, L. M., Aguaded, I., de Dios, S. C., Romero-Rodríguez, L. M., & Aguaded, I. (2020). El lenguaje como creador de realidades y opinión pública: Análisis crítico a la luz del actual ecosistema mediático. *Íconos. Revista de Ciencias Sociales*, (67), 139-157. <https://doi.org/10.17141/iconos.67.2020.3942>

Domínguez-Gómez, J. A. (2022). La modernización del análisis de contenido. Ejemplos de aplicación en evaluación socioambiental. *New Trends in Qualitative Research*. https://www.scielo.pt/scielo.php?script=sci_arttext&pid=S2184-77702022000500008&lang=en

Duoc, B. (2025). *Biblioteca: Investigación Aplicada, Innovación y Transferencia: Definición y propósito de la Investigación Aplicada*.

<https://bibliotecas.duoc.cl/investigacion-aplicada/definicion-proposito-investigacion-aplicada>

Estrada, J. C., Nacipucha, N. S., & Duarte, W. D. (2021). Marketing de contenidos y decisión de compra de los consumidores Generación Z en Ecuador. *ECA Sinergia*, 12(2), 25-37.

Flores, J., & Zamora, W. (2023). *Metodologías Cuantitativa y Cualitativa en la Investigación Científica: Un abordaje desde la epistemología – Revista Jireh UML*. <https://revistajireh.uml.edu.ni/publicaciones/vol-4-num-2-2024/24209-2/>

Galioto, M., Pedone, F., Vantarakis, A., Tavares, P., & Bianco, A. (2025). The use of social networks in institutional communication at university: A prospective pattern for enhancing the sense of belonging among students. *Frontiers in Communication*, 10. <https://doi.org/10.3389/fcomm.2025.1523295>

Guevara Albán, G. P., Verdesoto Arguello, A., & Castro Molina, N. E. (2020a). Metodologías de investigación educativa (descriptivas, experimentales, participativas, y de investigación-acción). *RECIMUNDO: Revista Científica de la Investigación y el Conocimiento*, 4(3), 163-173.

Guevara Albán, G. P., Verdesoto Arguello, A., & Castro Molina, N. E. (2020b). Metodologías de investigación educativa (descriptivas, experimentales, participativas, y de investigación-acción). *RECIMUNDO: Revista Científica de la Investigación y el Conocimiento*, 4(3), 163-173.

Guzmán, J. C. V., Domínguez, E. G., Reséndiz, P. Z., Clemente, A. P., & García, E. L. (2025). APIs RESTful para interoperabilidad entre aplicación Móvil y aplicación Web. *Ciencia Latina Revista Científica Multidisciplinar*, 9(3), 6795-6821. https://doi.org/10.37811/cl_rcm.v9i3.18325

Manrique-Gómez, L., Borja Gómez, J. H., Manrique-Gómez, L., & Borja Gómez, J. H. (2025). Ciencia de datos para la historia: Datificar las fuentes para una historia (predictiva). *Historia y grafía*, (64), 97-145. <https://doi.org/10.48102/hyg.vi64.541>

Manterola, C., Rivadeneira Dueñas, J., & Otzen Hernández, T. (2024). La Pregunta de Investigación y su Asociación con los Niveles de Evidencia y Grados de Recomendación en Estudios de Investigación Clínica. *International Journal of morphology*, 42(4), 1020-1028.

Mayet Comerón, T., Alonso Berenguer, I., Gorina Sánchez, A., Mayet Comerón, T., Alonso Berenguer, I., & Gorina Sánchez, A. (2022). Visualización de información y conocimiento para comunicar resultados de investigación de las ciencias sociales. *Mendive. Revista de Educación*, 20(3), 772-789.

Natarajan, T., & Pichai, S. (2024). Transition From Waterfall to Agile Methodology: An Action Research Study. *IEEE Access*, 12, 49341-49362. <https://doi.org/10.1109/ACCESS.2024.3384097>

Pérez Alvarez, D. M., Anías Calderón, C., Pérez Alvarez, D. M., & Anías Calderón, C. (2024). Integración de la gestión de redes empleando APIs. *Ingeniería Electrónica, Automática y Comunicaciones*, 45(1), 51-69.

Pressman, R. S., & Maxim, B. R. (2020). *Software engineering: A practitioner's approach* (Ninth edition). McGraw-Hill Education.

Sánchez, D. E. M., Vásquez, M. C. S., Morales-Urrutia, D., & Medina, R. P. (2025). Análisis cuantitativo de redes sociales en universidades públicas—Zona 3 Ecuador. *REVISTA MULTIDISCIPLINARIA DE DESARROLLO AGROPECUARIO, TECNOLÓGICO, EMPRESARIAL Y HUMANISTA.*, 7(2), 11-11.

Sánchez, D. M. V., Sánchez, M. J. P., & Alcívar, D. K. L. (2023). Improving speaking skills by implementing self-assessment and self-regulation facilitated by whatsapp in 8th year students. *Ciencia Latina Revista Científica Multidisciplinar*, 7(1), 2601-2619. https://doi.org/10.37811/cl_rcm.v7i1.4611

Sarmiento, C. F. H., Criollo, D. R. P., González, L. P. C., Ochoa, I. S. C., & Tamayo, L. A. C. (2024). Propuestas Analíticas para la Interpretación de las Obras Introducción Tema y Variaciones de Gioachino Rossini y Dance Preludes de Witold

Lutoslawski. *Ciencia Latina Revista Científica Multidisciplinar*, 8(1), 10092-10102. https://doi.org/10.37811/cl_rcm.v8i2.10325

Serrano-Orellana, K., Molina-Andrango, V., Franco-Gómez, M. del C., & Coronel-Ojeda, A. (2025). Uso de tecnologías digitales por grupos sociodemográficos en Ecuador: Análisis de la generación Z y los teléfonos inteligentes. *Sociedad & Tecnología*, 8(1), 18-39. <https://doi.org/10.51247/st.v8i1.459>

Terán, O., Narciso, F., Ríos-Bolívar, A., Hidrobo, F., Álvarez, J., León, L., Aguilar, J., & Hernández, D. (2009). Un marco metodológico para el desarrollo de aplicaciones para automatización industrial. *Revista de la Facultad de Ingeniería Universidad Central de Venezuela*, 24(1), 57-69.

Torrejón-Guirado, M.-C., San Martín Erice, I., San Martín Rodríguez, L., & Lima Serrano, M. (2024). Claves metodológicas y estratégicas para estudios basados en encuestas online: Un análisis basado en la iniciativa Checklist for Reporting Results of Internet E-Surveys. *Enfermería clínica*, 34(3), 207-213.

Valdiviezo, A., Cesibel, & Bonini, T. (2020). Uso de big data y data mining en los procesos de automatización de la comunicación de las organizaciones. *17 noviembre 2020*, 8(199), 128-142.

Vergara-Gonzalez, D. J., Villar-Ventura, L. D., Cano, M., Raymundo, C., & Dominguez, F. (2020). Método analítico de planificación operativa aplicando Servqual para la reducción de servicios no conformes en empresas de Saneamiento Ambiental: Decima Novena Conferencia Iberoamericana en Sistemas, Cibernética e Informatica, CISCi 2020, Decimo Septimo Simposium Iberoamericano en Educacion, Cibernética e Informatica, SIECI 2020 - 19th Ibero-American Conference on Systems, Cybernetics and Informatics, CISCi 2020, 17th Ibero-American Symposium on Education, Cybernetics and Informatics, SIECI 2020. *CISCi 2020 - Decima Novena Conferencia Iberoamericana en Sistemas, Cibernética e Informatica, Decimo Septimo Simposium Iberoamericano en Educacion, Cibernética e Informatica - Memorias, CISCi 2020 - Decima Novena Conferencia Iberoamericana en Sistemas, Cibernética e Informatica, Decimo Septimo Simposium Iberoamericano en Educacion, Cibernética e Informatica - Memorias*, 53-56.

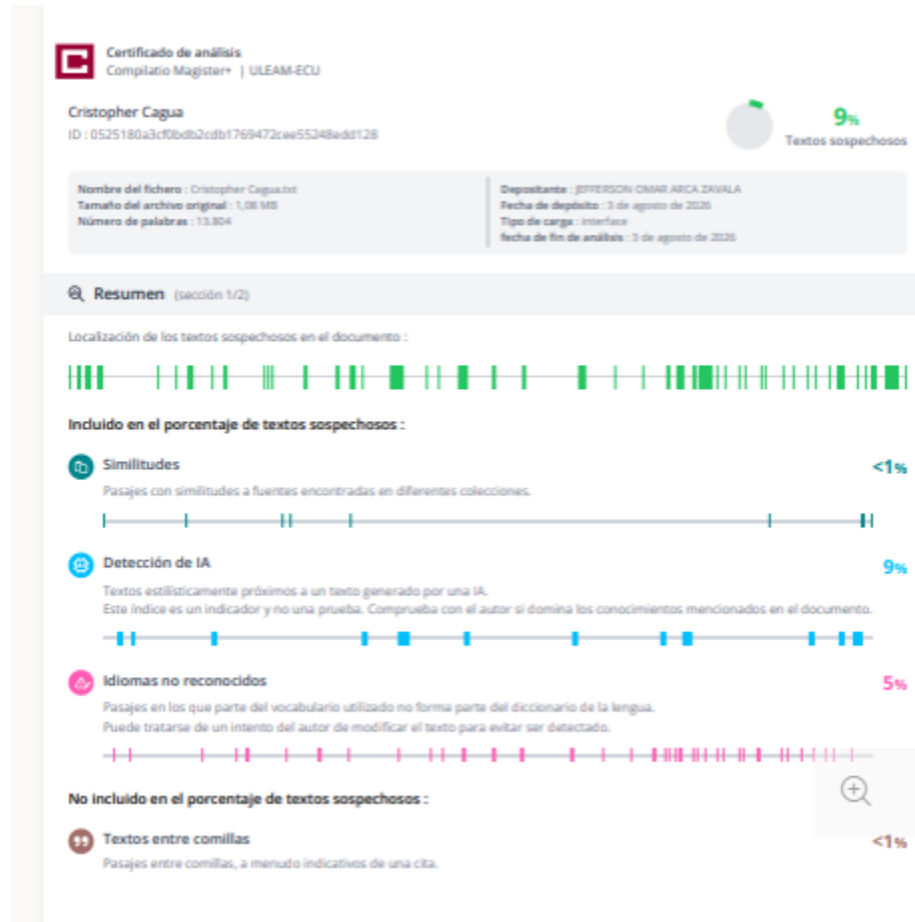
Villalobos-Murillo, J., Garita-González, G., Alfaro Ramírez, B. J., Villalobos-Murillo, J., Garita-González, G., & Alfaro Ramírez, B. J. (2025). Desarrollo de competencias: Inteligencia artificial y aprendizaje automático en prácticas supervisadas de estudiantes en computación. *Uniciencia*, 39(1), 32-50. <https://doi.org/10.15359/ru.39-1.3>

Yucra Quispe, T., & Bernedo Villalta, L. Z. (2020). Epistemología e Investigación Cuantitativa. *Igobernanza*, 3(12), 107-120.

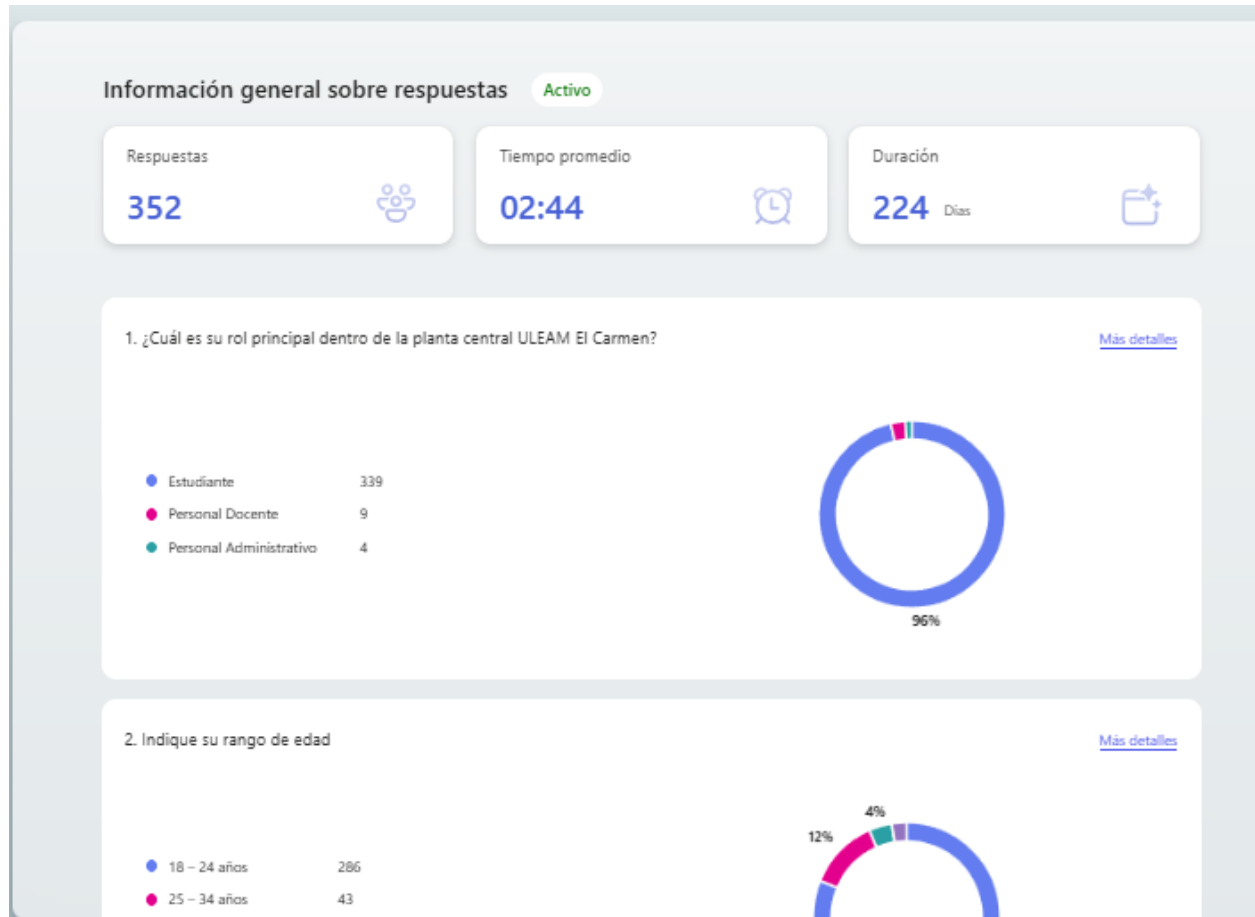
Zuasnábar, D. H., Kajatt, J. U., & Hidalgo, D. E. R.-C. (2023). Pagos digitales e inclusión financiera: Un estudio correlacional en microempresarios del distrito de Pueblo Libre - Lima, Perú 2022. *Industrial Data*, 26(2), 267-286.

ANEXOS

Anexo 1: Reporte del sistema anti plagio Compilatio: informe de similitud del trabajo donde muestra posibles coincidencias, su propósito, garantizar la originalidad del proyecto



Anexo 2: Evidencia de encuesta



Anexo 3: Evidencia de respuestas de estudiantes y personal administrativos

Id	Hora de inicio	Hora de finalización	Correo electrónico	Nombre	¿Cuál es su rol princ	Indique su rango de	¿Qué orientación vi	¿Cuál es el rango de	Al ver un
1	23/12/2025 0:31	23/12/2025 0:32	e1350732010@live.ule	ANDREA ROXANA DUEI	Estudiante	18 – 24 años	Vertical	Rango Estándar (20 a 6 Ambos	
2	23/12/2025 0:49	23/12/2025 0:51	e2300403918@live.ule	ANTHONY JOSUE BENA ANTHONY JOSUE BENA	18 – 24 años	Horizontal	Rango Estándar (20 a 6 Ambos		
3	23/12/2025 1:00	23/12/2025 1:02	e1315233864@live.ule	ARIANA ESTEFANIA ME	Estudiante	18 – 24 años	Vertical	Rango Micro (Menos di Ambos	
4	23/12/2025 1:03	23/12/2025 1:04	e2300434426@live.ule	AMELIA ESTERLINA VEF	Estudiante	18 – 24 años	Vertical	Rango Micro (Menos di Ambos	
5	23/12/2025 1:06	23/12/2025 1:07	e1728291806@live.ule	MARIA ANGELA SANTO	Estudiante	18 – 24 años	Vertical	Rango Micro (Menos di Solo auc	
6	23/12/2025 1:07	23/12/2025 1:09	e1316405412@live.ule	DANIELA LORELLY ERAZ	Estudiante	18 – 24 años	Horizontal	Rango Estándar (20 a 6 Ambos	
7	23/12/2025 5:47	23/12/2025 5:49	e2300575962@live.ule	DEYALI PAOLA MONTAL	Estudiante	18 – 24 años	Horizontal	Rango Estándar (20 a 6 Ambos	
8	23/12/2025 5:59	23/12/2025 6:00	e1317966156@live.ule	MELANIE DAYANA SAGI	Estudiante	18 – 24 años	Horizontal	Rango Micro (Menos di Ambos	
9	23/12/2025 6:13	23/12/2025 6:15	e1314707454@live.ule	ADRIANA LORENA CEDI	Estudiante	18 – 24 años	Vertical	Rango Estándar (20 a 6 Ambos	
10	23/12/2025 7:25	23/12/2025 7:26	e1314107796@live.ule	JHON ALEXANDER TOA	Estudiante	18 – 24 años	Horizontal	Rango Estándar (20 a 6 Solo sub	
11	23/12/2025 7:43	23/12/2025 7:44	e2300062284@live.ule	KRISTIE NATASHA HERR KRISTIE NATASHA HERR	18 – 24 años	Vertical	Rango Estándar (20 a 6 Ambos		
12	23/12/2025 7:43	23/12/2025 7:45	e0706283389@live.ule	DAVID JOSUE SALTOS N	Estudiante	18 – 24 años	Horizontal	Rango Documental (Mí Ambos	
13	23/12/2025 7:44	23/12/2025 7:46	e2300876550@live.ule	MYLADY LISSETH BRAV	Estudiante	18 – 24 años	Vertical	Rango Micro (Menos di Ambos	
14	23/12/2025 7:45	23/12/2025 7:47	e1313660753@live.ule	MARIA JOSE SOLORZAN	Estudiante	18 – 24 años	Horizontal	Rango Micro (Menos di Ambos	
15	23/12/2025 7:45	23/12/2025 7:47	e1729949253@live.ule	MAYRA GISELA MENEI	Estudiante	18 – 24 años	Horizontal	Rango Extendido (1 a 3 Ambos	
16	23/12/2025 7:47	23/12/2025 7:49	e1351893126@live.ule	ELIANA ANNABELLE VE	Estudiante	18 – 24 años	Vertical	Rango Estándar (20 a 6 Ambos	
17	23/12/2025 7:47	23/12/2025 7:49	e2300123425@live.ule	LISBETH IVONNE OJED/	Estudiante	18 – 24 años	Horizontal	Rango Estándar (20 a 6 Ambos	
18	23/12/2025 7:47	23/12/2025 7:50	e1315458214@live.ule	MARYURI MARLEI OLAI	Estudiante	18 – 24 años	Horizontal	Rango Extendido (1 a 3 Solo auc	
19	23/12/2025 7:47	23/12/2025 7:50	e1317461901@live.ule	KAREN SEIENA BARREI	Estudiante	18 – 24 años	Vertical	Rango Estándar (20 a 6 Ambos	

GLOSARIO

API: Conjunto estandarizado de definiciones, protocolos y reglas que permite la comunicación, transferencia e intercambio de información estructurada entre diferentes sistemas o aplicaciones independientes.

API RESTful: Estilo de arquitectura de software para sistemas distribuidos basado en el protocolo HTTP, que permite la consulta y manipulación de recursos mediante operaciones estandarizadas (GET, POST, PUT, DELETE) y respuestas formateadas en estructuras JSON.

Análisis de Sentimientos: Técnica especializada del Procesamiento de Lenguaje Natural orientada a procesar y clasificar automáticamente la polaridad o carga subjetiva de un texto, categorizándolo en valores positivos, negativos o neutros.

Aprendizaje Automático: Disciplina de la inteligencia artificial fundamentada en métodos estadísticos y modelos computacionales que facultan a los sistemas para identificar patrones y aprender de manera autónoma a partir de datos históricos sin requerir programación explícita regla por regla.

ASGI: Estándar de interfaz de servidor web para Python (implementado mediante Uvicorn) que permite la gestión eficiente de múltiples peticiones y operaciones concurrentes de forma no bloqueante o asíncrona.

Background Worker: Proceso o subproceso computacional que se ejecuta de forma autónoma e independiente de la interfaz de usuario, encargado de realizar tareas periódicas o de alta demanda de recursos, tales como la extracción automatizada de datos en redes sociales.

Broker de Mensajería (MQTT / Mosquitto): Servidor intermediario que gestiona la recepción, enrutamiento y entrega de paquetes de datos en tiempo real entre módulos emisores (mineros) y receptores (bases de datos o backend) bajo un patrón desacoplado de publicación y suscripción (Pub/Sub).

Contenerización (Docker): Tecnología de virtualización ligera a nivel de sistema operativo que empaqueta una aplicación, sus dependencias y configuraciones dentro de un contenedor aislado, garantizando un entorno de ejecución homogéneo y replicable.

Dashboard (Tablero de Control): Interfaz gráfica interactiva de visualización de datos diseñada para sintetizar, monitorizar y proyectar métricas masivas e indicadores clave de rendimiento (KPI) en tiempo real, facilitando la interpretación visual y la toma de decisiones.

Dataset (Conjunto de Datos): Colección estructurada de registros, mediciones o metadatos recopilados durante una investigación empírica o proceso de minería, organizados para su posterior análisis estadístico o algorítmico.

Desacoplamiento Arquitectónico: Principio de diseño de software en el cual los componentes o capas del sistema (presentación, lógica de negocio, procesamiento y almacenamiento) operan de manera independiente, minimizando las interdependencias mutuas y garantizando tolerancia a fallos.

Engagement (Compromiso / Interacción Digital): Indicador cuantitativo que mide el grado de interactividad y respuesta activa de una comunidad digital frente a una publicación (reacciones, me gusta, visualizaciones y comentarios).

FastAPI: Framework moderno de Python de alto rendimiento orientado a la construcción de APIs web, caracterizado por su soporte nativo de asincronía (async/await) y la validación estricta de esquemas de datos.

JSON (JavaScript Object Notation): Formato textual ligero, legible y universalmente utilizado para el intercambio e interoperabilidad de datos entre clientes web, servidores y servicios externos.

JWT (JSON Web Token): Estándar abierto y seguro (RFC 7519) utilizado para la autenticación y autorización en sistemas web, el cual transmite la identidad y privilegios del usuario mediante un token digital firmado y compacto.

KDD (Knowledge Discovery in Databases / Descubrimiento de Conocimiento en Bases de Datos): Proceso iterativo y multietapa orientado a transformar datos crudos en conocimiento útil y aplicable mediante fases sucesivas de selección, preprocesamiento, transformación, minería de datos e interpretación de patrones.

Microservicios: Enfoque arquitectónico de desarrollo de software donde una aplicación se estructura como un conjunto de servicios modulares, autónomos y comunicados mediante protocolos ligeros de red.

Minería de Datos (Data Mining): Etapa técnica central del proceso KDD que aplica algoritmos computacionales para la identificación de patrones, tendencias y correlaciones no evidentes dentro de grandes volúmenes de datos.

Minería de Texto (Text Mining): Rama de la minería de datos centrada en la extracción, depuración y estructuración de información significativa a partir de fuentes de texto no estructuradas (como hilos de comentarios o descripciones de publicaciones).

Modelo en Cascada (Waterfall): Metodología secuencial clásica para la gestión del ciclo de vida del software (SDLC), en la cual el avance del proyecto se ejecuta linealmente a través de fases estrictamente delimitadas: requisitos, diseño, implementación, pruebas y despliegue.

OAuth 2.0: Protocolo estándar de autorización que permite a una aplicación obtener acceso seguro y delegado a recursos de un usuario en un servicio de terceros (como Meta o TikTok) mediante tokens temporales y sin exponer credenciales privadas.

ORM (Object-Relational Mapping / Mapeo Objeto-Relacional): Técnica de programación (implementada mediante SQLAlchemy) que transforma las tablas y

restricciones de una base de datos relacional en modelos y clases de objetos dentro del lenguaje de desarrollo.

PostgreSQL: Sistema de gestión de bases de datos relacional y orientado a objetos (RDBMS) de código abierto, caracterizado por su estricto apego a las propiedades ACID (atomicidad, consistencia, aislamiento y durabilidad).

Procesamiento de Lenguaje Natural (PLN / NLP): Campo interdisciplinario de la inteligencia artificial y la lingüística computacional enfocado en posibilitar que los sistemas informáticos analicen, procesen y extraigan significado del lenguaje humano.

Pruebas de Caja Blanca: Técnica de verificación y aseguramiento de calidad de software donde se inspecciona la lógica interna, las coberturas de funciones, el manejo de excepciones y las estructuras de control del código fuente.

Pruebas de Caja Negra: Metodología de prueba funcional en la cual se evalúan las entradas y salidas de la aplicación desde la perspectiva del usuario final, sin examinar el código de programación interno.

Pruebas de Integración: Fase de evaluación orientada a comprobar que los diferentes módulos o servicios desacoplados del sistema (pipeline de datos, broker, backend y base de datos) interactúen entre sí de manera armónica y sin pérdidas de información.

React: Biblioteca de código abierto desarrollada en JavaScript orientada a la creación de interfaces de usuario dinámicas y reactivas basadas en la arquitectura de componentes.

SQLAlchemy: Librería y ORM para Python que proporciona herramientas completas para la gestión y abstracción de bases de datos relacionales, previniendo vulnerabilidades como la inyección SQL.

Tailwind CSS: Framework de diseño web basado en clases de utilidad que facilita la maquetación responsiva y estilizada de interfaces de usuario.

UML (Unified Modeling Language / Lenguaje Unificado de Modelado): Notación gráfica estandarizada en la ingeniería de software para modelar, estructurar y

documentar visualmente la arquitectura y el comportamiento de un sistema (casos de uso, secuencia, clases y estados).

UX/UI (User Experience / User Interface): Disciplinas de diseño orientadas a construir la disposición gráfica (UI) y garantizar una interacción accesible, intuitiva y satisfactoria del usuario final con el aplicativo (UX).