

Conductual explorer: una solución autónoma frente al desafío del entorno cognitivo

*Pablo Caneda González
María Porrar Santisteban
Álvaro Álvarez Álvarez*

Resumen

En el tiempo de hiperconectividad en el que vivimos, los medios por los cuales las personas se comunican son, principalmente, digitales y accesibles por cualquier usuario de la red. Esto pone a nuestro alcance un sinnúmero de opiniones y publicaciones subjetivas al alcance de la pantalla de nuestro dispositivo. En el ámbito cognitivo relativo a la influencia y la inteligencia, supone un reto el aprovechamiento de este flujo de información, ya que no son cantidades de datos procesables por uno o varios recursos humanos en un espacio de tiempo razonable.

A tal efecto, surge la iniciativa, dentro del ámbito de las FAS y, más concretamente, en el área de la influencia e inteligencia de nivel operacional, de desarrollar de manera autónoma, sin apoyo de agentes externos y con recursos limitados, una solución tecnológica a la que se ha bautizado como Conductual Explorer, que permita manejar vastas cantidades de datos obtenidas de medios de en formato digital o redes sociales de un entorno geográfico o social delimitado.

El análisis y almacenamiento de toda esta información, permite hacer un seguimiento del entorno informativo objetivo a lo largo del tiempo, pudiendo detectar tendencias y narrativas y su evolución en los periodos temporales deseados. El procesamiento de esta información será de gran utilidad en el apoyo a la toma de decisiones al nivel al que nos desempeñamos.

Palabras clave

Redes sociales, Fuentes abiertas, Influencia, Inteligencia, Inteligencia artificial.

Behavioral explorer: an autonomous solution to the challenge of the cognitive environment

Abstract

In today's era of hyperconnectivity, digital communication channels are widely accessible to any internet user, exposing individuals to an overwhelming volume of subjective opinions and publications, instantly available at the touch of a screen. In the cognitive domain of influence and intelligence, leveraging this vast flow of information presents a significant challenge, as the sheer volume of data surpasses the processing capacity of human analysts within a reasonable timeframe.

To address this challenge, within the framework of the Armed Forces—specifically in the field of operational-level influence and intelligence—a fully autonomous technological solution has been developed without external support and with limited resources. This solution, known as Conductual Explorer, facilitates the processing of vast amounts of data obtained from digital and social media within a specified geographical or social context.

By analyzing and storing this information, the system enables continuous monitoring of the targeted informational landscape over time, allowing for the identification of emerging trends, narratives, and their evolution across selected timeframes. The insights derived from this process provide valuable decision-making support at the operational level.

Keywords

Social media, Open sources, Influence, Intelligence, artificial intelligence.

1 Introducción

«Al tomar posesión de su dignidad, César fue el primero que estableció que se llevaría un diario de todos los actos del senado y del pueblo, y que este diario se haría público».

Según cuenta Suetonio, en el año 59 a. C., Julio Cesar da la orden de publicar las actas diurnas (*Acta diurna populi Romani*), expuestas de forma periódica en lugares públicos, tenían la finalidad de proporcionar a la ciudadanía romana un canal de información más preciso que los chismes y cotilleos que viajaban de boca en boca por las intrincadas redes de la compleja vida social romana.

Consideradas por muchos como el primer medio de comunicación escrito y, por otros, como la primera publicación oficial a modo de primigenio boletín oficial del Estado, pronto se consolidaron como una herramienta para influir en la opinión pública. Como afirma el emperador Augusto, sucesor de César, es innegable su utilidad como herramienta de propaganda al servicio del Estado.

Desde entonces, la relevancia de la capacidad de influir en audiencias ha crecido paralelamente a la complejidad y tamaño de las organizaciones humanas y de forma más acusada en la actualidad, con la irrupción de nuevas tecnologías.

Hoy en día, los medios de comunicación tradicionales se integran con nuevas formas de difusión de información, nuevas aplicaciones y redes sociales permiten a millones de usuarios participar en la generación y difusión de contenidos de forma descentralizada. Esta ingente producción unida a la hiperconectividad generada en entornos digitales, es capaz de producir o modificar rápidamente corrientes de opinión que deben ser tenidas muy en cuenta en el estudio de la nueva meta-geopolítica.

En el actual contexto de conflictos híbridos, la «zona gris» se ha consolidado como un espacio de disputa estratégica donde las narrativas, la desinformación y la manipulación de la opinión pública desempeñan un papel crucial. La proliferación de plataformas digitales ha transformado el entorno informativo en un campo de batalla no convencional, en el que actores estatales y no estatales emplean tácticas de guerra cognitiva para influir en la percepción de la realidad. En este escenario, la capacidad para analizar el sentimiento en redes sociales (RRSS) y medios de comunicación (MMCC), así como clasificar grandes volúmenes de datos de manera eficiente, se convierte en un recurso clave para la defensa y la seguridad nacional.

Según nuestra doctrina, la inteligencia engloba una serie de actividades que permiten obtener una visión integral y actualizada del entorno donde

se desarrollan las operaciones, la necesidad de obtener una visión integral del multidominio, obliga a contemplar dentro de las capacidades de inteligencia herramientas de análisis de nuevos ámbitos como, por ejemplo, el pujante dominio cognitivo.

El estudio y conocimiento de las condiciones socioculturales de determinado entorno permitirá encontrar indicadores en el dominio cognitivo que servirán de alerta ante posibles crisis o amenazas a la seguridad.

La inteligencia militar, en posesión de su dignidad, debe innovar y adaptarse a un nuevo contexto, como ya hiciera Julio Cesar al establecer las actas diurnas en la antigua Roma.

En el ámbito de la defensa, son ampliamente conocidas las enormes posibilidades de la observación en fuentes abiertas, que nos puede permitir obtener con rapidez la información necesaria para producir inteligencia básica sobre cualquier parte del mundo y puede servir de apoyo en la comprensión profunda del entorno operativo, así como para detectar tendencias y corrientes de pensamiento en una parte o un todo de una sociedad.

La actualización de nuevas capacidades lleva aparejados algunos desafíos. El gran volumen de información que se debe manejar supone un desafío para los analistas, la necesidad de filtrar información incorrecta, desinformación intencionada o tendenciosa requiere personal experimentado y con un adiestramiento específico, lo que se traduce en muchos recursos humanos y de tiempo.

Para paliar estos inconvenientes, se convierte en necesaria una estrecha colaboración entre el analista, conocedor del producto necesario para determinar tendencias y corrientes en un entorno social y el experto en tratamiento de datos, conocedor de técnicas que permiten obtener, almacenar y procesar grandes cantidades de datos.

Durante el desarrollo de Conductual Explorer se han tenido en cuenta tanto aspectos técnicos como necesidades y puntos de vista del analista, tratando de adaptar al máximo el potencial de la herramienta a las necesidades de análisis.

1.1 Motivación: resolver necesidades internas con creatividad y bajo coste

En el caso concreto que se va a desarrollar en este artículo, para hacer frente a este desafío de procesar grandes cantidades de información de fuentes abiertas, surge la necesidad de desarrollar herramientas que permitan interpretar la dinámica del entorno informacional de manera sistemática y eficiente. Los métodos tradicionales de análisis presentan limitaciones en términos de escalabilidad y rapidez, lo que dificulta la obtención de

inteligencia procesable en tiempo real. La solución que aquí se presenta responde a esta problemática mediante un enfoque innovador y de bajo coste, basado en la automatización y el uso de inteligencia artificial para la extracción y clasificación de datos y para la extracción de estadísticas y métricas ilustrativas del análisis del entorno informativo que se está realizando.

Este código, desarrollado íntegramente en lenguaje de programación Python, se ha desarrollado con el enfoque de ser una herramienta OSINT (Open-Source Intelligence). Ha sido diseñado para facilitar la labor de analistas e investigadores en el ámbito de la seguridad y la defensa. Su propósito es ofrecer una alternativa accesible y adaptable a los requerimientos específicos de diferentes contextos operativos, sin depender de infraestructuras costosas o soluciones comerciales privativas. Al combinar técnicas avanzadas de procesamiento del lenguaje natural (PLN), modelos de inteligencia artificial de *machine learning* y *deep learning*, así como técnicas de programación tradicional, la herramienta permite no solo analizar el sentimiento predominante en el entorno digital, sino también estructurar y categorizar grandes volúmenes de información provenientes de RRSS y MMCC con un alto grado de precisión y exponer los resultados de una manera fácilmente comprensible. Al mismo tiempo, es un ejemplo de aplicación de tecnologías programáticas avanzadas como solución a un problema concreto, empleando el recurso técnico adecuado para resolver cada arista del problema. Esto último pone en valor la idea de que las soluciones programáticas son un medio y no un fin y que no siempre la solución más moderna es la más adecuada para afrontar un problema.

1.2 Contexto del proyecto y su relevancia

El uso de información digital en escenarios de zona gris ha demostrado ser un factor determinante en la configuración de estrategias militares, diplomáticas y de seguridad. El conflicto en zona gris se caracteriza por la ausencia de confrontación directa, siendo el dominio de la información un elemento central en la lucha por la influencia y la legitimidad. En este contexto, la capacidad de anticipar tendencias, detectar patrones de desinformación y analizar discursos en tiempo real permite a los actores gubernamentales y privados tomar decisiones fundamentadas y estratégicas.

Este proyecto nace de la necesidad de contar con una herramienta flexible y eficiente que permita a los analistas gestionar el creciente volumen de datos en el ciberespacio. Su implementación ofrece ventajas en la detección temprana de amenazas, el monitoreo de campañas de influencia y la evaluación del impacto de narrativas en distintos públicos objetivos. Además, su carácter modular permite su integración con otras herramientas OSINT, optimizando así el proceso de obtención de inteligencia en entornos de alta volatilidad y riesgo.

Con esta iniciativa, se busca fortalecer la capacidad analítica en el ámbito de la seguridad, proporcionando una solución efectiva y accesible para el estudio del entorno informativo digital en zonas de conflicto.

1.3 Objetivos y alcance del programa

El principal objetivo de este programa es proporcionar una herramienta OSINT capaz de recopilar, analizar y clasificar datos provenientes de RRSS y MMCC en zonas de conflicto o que sean de interés, con el fin de optimizar la toma de decisiones estratégicas en materia de seguridad y defensa. Para ello, la herramienta se basa en técnicas de PLN, permitiendo una evaluación automatizada del sentimiento predominante en el entorno digital y una segmentación eficiente de los datos según criterios específicos.

Objetivos específicos

- Automatización del análisis de sentimiento: implementar modelos de PLN para identificar y categorizar emociones y tendencias en el discurso digital.
- Clasificación avanzada de datos: desarrollar modelos de inteligencia artificial que permitan segmentar información de RRSS y MMCC en función de su relevancia y confiabilidad.
- Presentar los resultados de una manera comprensible, con estadísticas y métricas que permitan entender de manera sencilla los resultados del análisis y extraer conclusiones.
- Monitoreo en tiempo real: diseñar una estructura de actualización continua para detectar cambios en la narrativa digital y evaluar su impacto.
- Adaptabilidad y escalabilidad: crear un sistema modular que pueda ajustarse a distintos entornos operativos y requerimientos específicos.
- Integración con otras herramientas OSINT: facilitar la interoperabilidad con bases de datos y sistemas de análisis ya existentes para maximizar la eficiencia operativa.

Alcance del programa

El programa se ha diseñado con un enfoque flexible y modular, lo que permite su implementación en distintos escenarios y niveles de operación. Está orientado tanto a entidades gubernamentales como a organizaciones privadas interesadas en el monitoreo de la información digital en contextos de crisis o conflicto.

El sistema podrá ser utilizado para:

- Análisis de desinformación y propaganda: identificar campañas de influencia y estrategias de manipulación informativa en plataformas digitales.
- Evaluación del impacto de narrativas: determinar cómo ciertos discursos afectan la percepción pública y el comportamiento de la audiencia.
- Apoyo en la toma de decisiones estratégicas: proporcionar datos estructurados y análisis predictivos para mejorar la respuesta a amenazas emergentes.
- Seguimiento de dinámicas de conflicto: monitorear el desarrollo de crisis en tiempo real mediante el análisis de contenido digital.

Este enfoque integral busca dotar a los analistas de inteligencia de una herramienta que no solo simplifique la gestión de grandes volúmenes de información, sino que también proporciona capacidades avanzadas de análisis predictivo y estratégico para la comprensión del entorno informativo deseado.

2 Necesidad de un analista y creación de la base de datos

El uso de herramientas de inteligencia de fuentes abiertas (OSINT) ha adquirido una relevancia creciente en el ámbito del análisis de datos, proporcionando un apoyo significativo al analista, sin sustituir la labor esencial del factor humano. Si bien la automatización permite procesar grandes volúmenes de información con rapidez y eficiencia, la interpretación contextual y la definición de criterios de búsqueda siguen requiriendo la intervención humana para garantizar la precisión y fiabilidad de los datos recopilados.

Antes de la implementación de una herramienta OSINT, es necesario establecer un marco metodológico que garantice la calidad del análisis. La fase inicial implica delimitar el entorno informativo, definir el objetivo de búsqueda y estructurar la información a explotar en una base de datos robusta. La ausencia de alguno o varios de estos pasos previos incrementa el margen de error y puede derivar en la generación de resultados incompletos o imprecisos.

En este capítulo, se detallan los procedimientos esenciales para la recopilación y estructuración de datos, destacando la importancia del analista en la construcción de una base de datos que facilite el procesamiento de información y la obtención de resultados válidos. Además, se aportan ejemplos con casos concretos para cada etapa del procedimiento.

2.1 Análisis de la Zona: recopilación y estructuración de datos

La delimitación del entorno de información constituye el primer paso en el análisis OSINT, debiendo establecerse criterios claros para la recopilación de datos. Esta fase comprende la identificación de parámetros espaciales y temporales, la selección de fuentes y audiencias relevantes y la identificación de actores clave dentro del ecosistema digital objetivo.

La metodología de recopilación de datos se organiza en tres etapas fundamentales:

2.1.1 Definición del ámbito de estudio

La delimitación del análisis debe atender a una serie de factores clave que garanticen la precisión y relevancia de la información recopilada. En primer lugar, resulta fundamental establecer los parámetros espaciales y temporales, definiendo con claridad el objeto de estudio y el periodo de observación, lo que permite contextualizar de manera adecuada los datos obtenidos.

Asimismo, es necesario determinar el sector de interés, identificando el ámbito temático sobre el cual se desarrollará el análisis. Este puede situarse en dimensiones políticas, sociales, económicas o de seguridad, dependiendo del objetivo de la investigación.

Por último, la selección de plataformas y medios de comunicación relevantes constituye un aspecto esencial del proceso. La identificación de fuentes primarias y secundarias, tales como redes sociales, foros, blogs y medios digitales especializados, permite la recopilación de información representativa del ecosistema informativo analizado.

2.1.2 Selección y filtrado de fuentes de información

El correcto funcionamiento de la herramienta OSINT desarrollada depende en gran medida de la precisión y fiabilidad de los datos que alimentan su base de análisis. Aunque la recopilación de información es un proceso complejo, su eficacia radica en la labor previa del analista, quien debe definir criterios claros de selección y garantizar que las fuentes sean verificables y representativas del entorno informativo. Sin esta fase inicial, el margen de error se incrementa, afectando directamente la capacidad del programa para procesar y clasificar datos de manera precisa, ofreciendo unos resultados que pueden diferir de la realidad.

Para optimizar su rendimiento, la herramienta OSINT integra datos provenientes de medios de comunicación, redes sociales y fuentes abiertas, priorizando aquellas que cumplen con estándares de fiabilidad. Además,

emplea filtros automáticos y técnicas de normalización para minimizar ruido informativo, detectar patrones de comunicación y estructurar la información de manera coherente. En el caso de redes sociales, se analizan tendencias discursivas, uso de símbolos y *hashtags*, y, cuando es posible, datos geolocalizados para mejorar la contextualización.

La flexibilidad en la selección de datos es clave para adaptar la herramienta a distintos escenarios, permitiendo ajustar parámetros en función de la disponibilidad y calidad de la información. En este sentido, la labor del analista no solo define el punto de partida del proceso, sino que también optimiza el rendimiento del programa, asegurando que las predicciones y clasificaciones generadas sean lo más precisas y útiles posible.

2.1.3 Identificación de actores clave

Dentro del ecosistema digital, ciertos actores poseen un papel determinante en la difusión y amplificación de la información, influyendo en la construcción de narrativas y en la percepción pública de los acontecimientos. La identificación de estos agentes es un paso esencial en el análisis OSINT, ya que permite comprender cómo se propagan los discursos, quiénes los generan y qué impacto tienen dentro del entorno informativo.

Para ello, la herramienta OSINT implementa un proceso de segmentación de usuarios por nivel de influencia, diferenciando entre actores gubernamentales, medios de comunicación, líderes de opinión y comunidades digitales. Esta clasificación facilita la detección de fuentes clave dentro de cada categoría y permite analizar su rol en la difusión de la información, ya sea como emisores primarios, amplificadores o agentes de contranarrativa.

Además, se lleva a cabo un análisis de interacción y alcance, que evalúa la capacidad de cada actor para generar reacciones dentro del ecosistema digital. Este proceso considera métricas como la frecuencia de publicación, el número de interacciones (comentarios, compartidos, reacciones) y el alcance del contenido. A través de estos indicadores, la herramienta puede identificar patrones de difusión y detectar tendencias emergentes, lo que resulta fundamental para anticipar el impacto de ciertos discursos y evaluar su nivel de penetración en distintos segmentos de la audiencia.

La combinación de estas metodologías permite a la herramienta OSINT no solo mapear el flujo de información dentro del ecosistema digital, sino también optimizar la precisión del análisis, distinguiendo entre actores con verdadera capacidad de influencia y aquellos cuya presencia es circunstancial o de menor relevancia.

Ejemplo aplicado: análisis de la percepción de la UME y el Ejército español en la Comunidad Valenciana tras la DANA

En este caso, el analista debe recopilar datos sobre cómo se ha percibido la intervención de la UME y del Ejército Español tras la crisis de la DANA.

Delimitación geográfica y sectorial

Para garantizar la precisión del análisis, se ha delimitado la zona geográfica y el ámbito de estudio en función de la zona de operación y la temática de interés. En este caso, la investigación se centra en la Comunidad Valenciana, con un enfoque en la percepción mediática y social de la intervención de las fuerzas militares, diferenciando la cobertura y el discurso en torno a la Unidad Militar de Emergencias (UME) y el Ejército español en el contexto de la DANA.

Fuentes de Información

La recopilación de datos se ha realizado a partir de fuentes digitales representativas del ecosistema informativo valenciano. Se han seleccionado los siguientes canales para garantizar una visión integral del discurso mediático y ciudadano:

- Medios digitales: se han analizado artículos y reportajes publicados en *Levante-EMV*, *Las Provincias*, *À Punt* e *Información Alicante*, debido a su relevancia en la cobertura informativa regional.
- Red social X: se han recopilado publicaciones de ciudadanos, periodistas y actores políticos con el fin de evaluar la conversación digital y su impacto en la percepción pública.
- Telegram: se han extraído publicaciones de cuentas clave con el ánimo de analizar y clasificar sus mensajes.
- Plataformas comunitarias: se han examinado debates y opiniones en grupos de Facebook y foros locales, permitiendo identificar la percepción de la ciudadanía fuera de los canales mediáticos convencionales.

Filtrado y Normalización de Datos

Dado el gran volumen de información recopilada, ha sido necesario implementar un proceso de filtrado y normalización de datos con el objetivo de garantizar la pertinencia y la calidad del análisis. Este proceso ha comprendido las siguientes etapas:

- Eliminación de publicaciones irrelevantes: se han descartado contenidos que no guardan relación directa con la DANA o con la intervención de las fuerzas militares, reduciendo el ruido informativo.

- Normalización de datos: se ha realizado la estandarización de fechas, formatos y categorías, asegurando la coherencia de la información y su compatibilidad con la herramienta OSINT desarrollada.

2.2 Herramientas empleadas en el análisis inicial

Una vez que el analista ha realizado una búsqueda precisa de fuentes y actores clave, estos se introducen en herramientas especializadas para la extracción automatizada de datos, optimizando el proceso de recopilación. Para ello, se emplean soluciones comerciales y de acceso abierto que permiten obtener información estructurada de diversas plataformas.

Entre estas herramientas se incluyen técnicas de *web scraping*, que permiten extraer contenido de sitios web de interés, APIs de redes sociales, que facilitan la recopilación de publicaciones y metadatos de plataformas como Telegram o TikTok y plataformas de agregación de noticias, que centralizan información proveniente de diferentes medios.

Si bien la construcción de la base de datos se realiza manualmente, el analista puede apoyarse en hojas de cálculo avanzadas, *software* de gestión de datos y sistemas de etiquetado automático, lo que facilita la organización y segmentación de la información recopilada, asegurando su coherencia y utilidad para el análisis posterior.

Ejemplo aplicado: uso de herramientas para analizar la percepción de la UME y el Ejército español

En este caso, una vez que el analista ha identificado las fuentes clave, se emplean herramientas especializadas para la extracción de datos y su posterior análisis. Para ello, se pueden utilizar:

- Técnicas de *web scraping*: se han aplicado métodos de extracción automatizada de datos para recopilar artículos de medios digitales sobre la DANA en la Comunidad Valenciana, permitiendo un análisis exhaustivo de la cobertura informativa.
- APIs de redes sociales: se han utilizado interfaces de programación de aplicaciones (APIs) para identificar menciones y patrones discursivos en plataformas como Telegram, con el objetivo de analizar la conversación digital en torno a la UME y el Ejército español. Esta metodología permite acceder a publicaciones en tiempo real y evaluar la interacción de los usuarios con determinados discursos.
- Plataformas de agregación de noticias: se han empleado herramientas especializadas que centralizan información de diversos medios, facilitando la detección de tendencias informativas y la identificación de narrativas predominantes en el contexto de la DANA.

2.3 Construcción de una base de datos

Una vez recopilada la información, es fundamental estructurarla y consolidarla en una base de datos robusta que garantice su almacenamiento seguro y facilite su explotación eficiente en el análisis OSINT. La calidad y organización de esta base de datos determinarán, en gran medida, la precisión de los resultados obtenidos por la herramienta, ya que un cribado deficiente puede derivar en errores de interpretación, sesgos o pérdida de información relevante.

Para optimizar su funcionamiento, la base de datos debe cumplir con una serie de criterios de organización que permitan una gestión eficaz de los datos recopilados. En primer lugar, la estructura de almacenamiento debe contemplar una categorización detallada basada en fecha, fuente, tipo de contenido y nivel de relevancia. Esto permite establecer filtros que agilicen el acceso a la información y faciliten la identificación de patrones dentro del conjunto de datos.

El etiquetado de información es otro aspecto clave en la optimización del almacenamiento. El uso de palabras clave, *hashtags* y categorías contribuye a mejorar la segmentación y búsqueda de datos dentro del repositorio de almacenamiento, permitiendo una indexación más precisa de los elementos disponibles. La integración de múltiples formatos es también un requisito indispensable, dado que la información recopilada puede presentarse en diversos tipos de archivos, incluyendo texto, imágenes, videos y archivos de audio. La capacidad de almacenar y procesar estos formatos garantiza que la herramienta pueda analizar datos heterogéneos sin limitaciones estructurales.

Además de la organización, la seguridad y mantenimiento de la base de datos es un factor crítico para garantizar su fiabilidad y continuidad operativa. La protección de datos es prioritaria, por lo que se implementan medidas para prevenir la manipulación de la información y evitar pérdidas accidentales. Esto implica la aplicación de protocolos de autenticación, control de acceso y auditorías periódicas para verificar la integridad de los registros.

Asimismo, la base de datos debe contar con un sistema de actualización continua, que permita la incorporación de nuevas fuentes de información y la eliminación de datos obsoletos. Este proceso asegura que el análisis realizado por la herramienta OSINT se base siempre en información actualizada y relevante, evitando sesgos provocados por el uso de datos desactualizados o irrelevantes.

En conjunto, estos criterios permiten la construcción de una base de datos estructurada, segura y adaptable, garantizando que la herramienta pueda procesar información de manera eficiente y ofrecer análisis precisos basados en datos de calidad.

Por último, almacenar de manera estructurada los datos recopilados de fuentes abiertas, así como los resultados obtenidos de los análisis y clasificaciones realizadas por el programa o el analista, será fundamental a la hora de generar o refinar modelos de IA, ya que estos modelos requieren de cantidades considerables de datos etiquetados. Estos modelos realizan tareas fundamentales a la hora de llevar a cabo los análisis, como se detallará más adelante.

Ejemplo aplicado: construcción de la base de datos sobre la percepción de la UME y el Ejército español en la Comunidad Valenciana

El analista debe diseñar una base de datos estructurada que garantice la organización eficiente de la información y facilite su explotación en fases posteriores del análisis OSINT. Para ello, la base de datos debe incluir:

- Relación de usuarios clave de redes sociales y medios de comunicación relevantes para el análisis.
- Registro de menciones en medios de comunicación, incorporando datos clave como la fecha de publicación, la fuente y la categoría de información, lo que permitirá evaluar la cobertura mediática y su evolución en el tiempo.
- Análisis de publicaciones en redes sociales, con una recopilación sistemática de contenido asociado a hashtags relevantes, palabras y emoticonos utilizados o métricas de interacción, posibilitando la identificación de tendencias y el impacto de determinados discursos en la conversación digital.
- Categorización por percepción pública, segmentando el contenido según su tono positivo, negativo o neutro, con el objetivo de evaluar la orientación predominante en la narrativa mediática y social.

Una base de datos densamente poblada y bien estructurada es fundamental para la implementación de herramientas analíticas avanzadas, ya que permite la detección de patrones de opinión, el análisis de tendencias y la evaluación del impacto de narrativas específicas en el ecosistema informativo.

3 Códigos desarrollados y su funcionamiento

3.1 Código 1: análisis de sentimiento del entorno informativo

3.1.1 Objetivo y lógica detrás del análisis

El análisis de sentimiento es una técnica fundamental en el procesamiento de información en entornos digitales, utilizada para evaluar la percepción y actitud expresada en publicaciones de redes sociales (RRSS) y

medios de comunicación (MMCC). Su aplicación en la herramienta OSINT desarrollada tiene como objetivo principal determinar la orientación emocional y narrativa del contenido analizado, proporcionando una visión estructurada sobre cómo se perciben determinados eventos, actores o temas dentro de un ecosistema informativo.

El propósito del análisis de sentimiento en este contexto es clasificar y cuantificar el tono predominante en las publicaciones asociadas a una operación o estudio específico. Para ello, se emplean modelos avanzados de PLN que asignan valores a las publicaciones en función de su carga emocional, segmentándolas en tres categorías principales:

- Positivas: contenido con una percepción favorable hacia el tema analizado.
- Negativas: publicaciones con un tono crítico, adverso o hostil.
- Neutras: información sin carga emocional evidente o con un tono equilibrado.

La lógica subyacente de este análisis radica en la necesidad de comprender la dinámica del discurso digital, identificando tendencias, patrones de polarización y posibles intentos de manipulación informativa. Gracias a esta metodología, es posible detectar cambios en la percepción pública, evaluar el impacto de ciertas narrativas y establecer correlaciones entre eventos específicos y la reacción del ecosistema digital.

Además, este enfoque no solo permite la monitorización en tiempo real, sino que también aporta información estratégica para la toma de decisiones, facilitando la adaptación de mensajes, la identificación de riesgos comunicacionales y el desarrollo de estrategias de respuesta basadas en datos objetivos.

3.1.2 Técnicas utilizadas: procesamiento del lenguaje natural (PLN) y clasificación por taxonomías

El análisis de sentimiento implementado en este código se basa en técnicas avanzadas de PLN, permitiendo que los algoritmos comprendan, procesen y clasifiquen textos en lenguaje natural de manera estructurada y precisa. Se ha utilizado, en principio, un modelo de *deep learning* entrenado para determinar el sentimiento de un texto.

Para llevar a cabo este análisis, primero se realiza una fase de preprocesamiento del texto, donde se eliminan caracteres especiales, emojis y signos de puntuación que podrían generar ruido en los resultados. Además, el texto se convierte a minúsculas para evitar discrepancias debidas a la capitalización y se eliminan las palabras vacías (*stopwords*), aquellas que no aportan información relevante, como artículos y conjunciones.

Una vez limpiado el texto, este debe transformarse en una representación numérica comprensible para los modelos de análisis. Para ello, se emplean técnicas como TF-IDF (Term Frequency-Inverse Document Frequency), que pondera la relevancia de las palabras en un conjunto de textos.

Para la fase de clasificación del sentimiento, se procesa el texto mediante un modelo entrenado para clasificación de sentimiento con base en el modelo BERT de clasificación de textos. BERT es un modelo de *deep learning* el cual se ha afinado para la clasificación deseada en este caso. De esta manera, se consigue asignar una categoría al texto procesado de cada publicación: positivo, cuando el contenido transmite una percepción favorable; negativo, cuando el tono es crítico o adverso; y neutro, cuando no hay una carga emocional predominante.

Los resultados del análisis se presentan mediante diversas representaciones visuales que facilitan la interpretación de los datos. Se utilizan gráficos de barras para mostrar el volumen de publicaciones en cada categoría, diagramas circulares que reflejan la proporción relativa de cada tipo de sentimiento y mapas de calor que posicionan a los actores clave dentro del ecosistema informativo en función de su favorabilidad u hostilidad, así como de su nivel de relevancia.

3.1.3 Ejemplo práctico y resultados esperados

El sistema desarrollado clasifica todas las publicaciones y noticias en función de criterios predefinidos por el analista, que tienen gran peso a la hora de entrenar el modelo de clasificación de texto utilizado, determinando el grado de favorabilidad o desfavorabilidad hacia el objetivo de la operación que esté siendo objeto de análisis.

Análisis de sentimientos en publicaciones

En la primera imagen proporcionada, se observa la distribución del sentimiento de las publicaciones analizadas. En este gráfico se presentan dos representaciones:

- Diagrama de barras: muestra la cantidad absoluta de publicaciones positivas, negativas y neutras. En la gráfica, se observa que la mayoría de las publicaciones son neutras, seguidas por las negativas y, en menor medida, las positivas.
- Gráfico circular: expresa la proporción porcentual de cada categoría. En este caso, se visualiza que el 77,4 % del contenido es neutro, el 13,2 % negativo y solo el 9,4 % positivo.

En la segunda figura, se presenta un gráfico de barras apiladas, donde se analizan palabras clave específicas determinadas por el analista, que han sido encontradas en las publicaciones mediante algoritmos tradicionales de programación y su asociación con un sentimiento determinado. En este análisis se destacan:

- El volumen de menciones de cada término en las publicaciones analizadas.
- La proporción de sentimiento (positivo, negativo o neutro) de cada palabra clave.

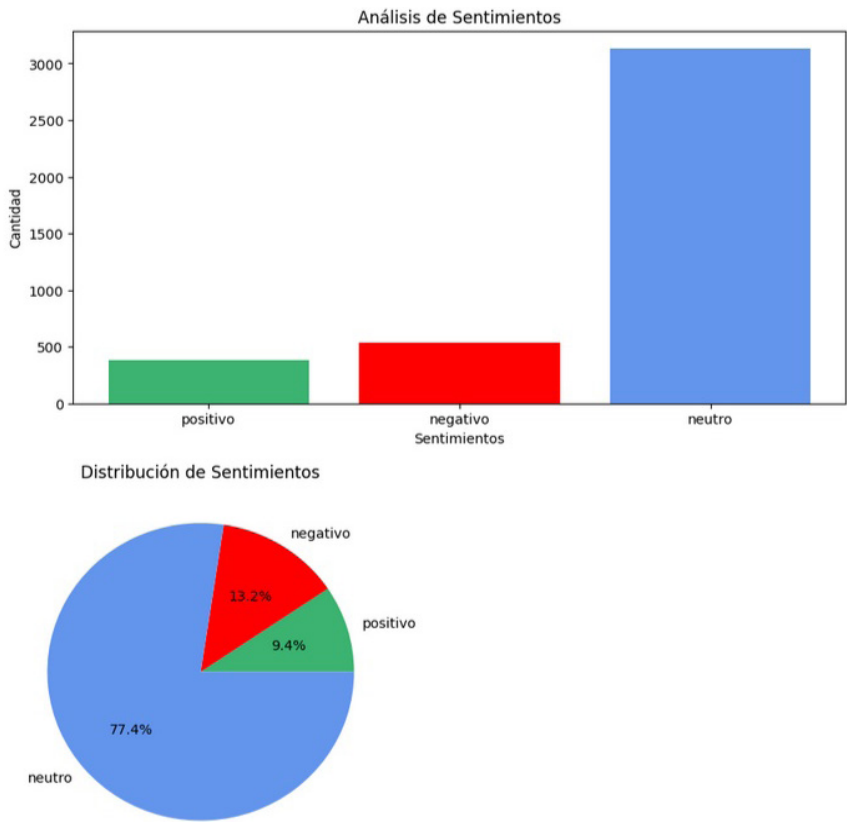


Figura 1. Gráficos de resultados de análisis de sentimiento. Elaboración propia

Este tipo de análisis permite identificar términos con una alta carga emocional negativa, los cuales pueden señalar la presencia de conflictos o controversias. Asimismo, facilita la evaluación de la percepción de conceptos clave dentro del ecosistema informativo, permitiendo un análisis más profundo de su impacto. Además, contribuye a la detección de oportunidades

estratégicas para influir en la narrativa digital a través de acciones comunicacionales dirigidas.

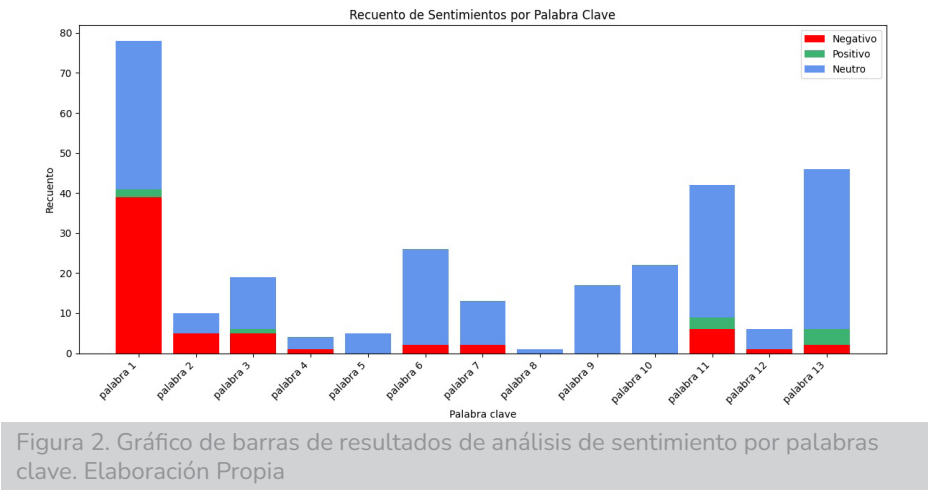


Figura 2. Gráfico de barras de resultados de análisis de sentimiento por palabras clave. Elaboración Propia

Por último, el mapa de calor (*heatmap*) utilizado en este análisis posiciona a los actores clave en función de dos variables fundamentales: su hostilidad o favorabilidad hacia el tema de interés objeto del análisis, representado según su posición en el eje x, y su relevancia dentro del entorno informativo, representado según la posición en el eje y. Esta representación gráfica permite identificar patrones de polarización y priorizar aquellos actores con mayor impacto en la conversación digital. Cada medio viene representado por una etiqueta con un número.

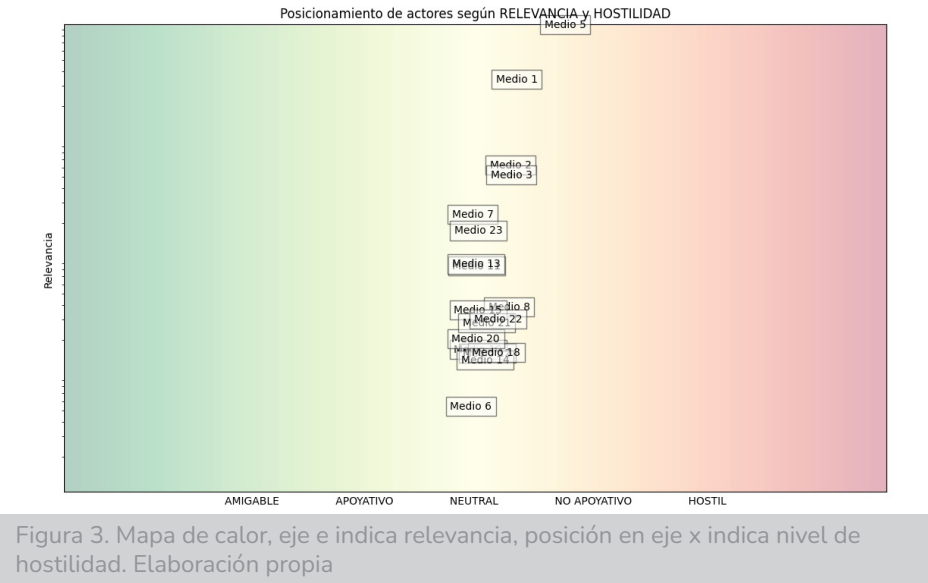


Figura 3. Mapa de calor, eje e indica relevancia, posición en eje x indica nivel de hostilidad. Elaboración propia

Para evitar que las diferencias de relevancia entre medios de comunicación y usuarios de redes sociales distorsionen la visualización, el eje y, que representa este valor, se encuentra en escala logarítmica. Dado que en entornos digitales la distribución de la relevancia suele ser altamente desigual —con algunos actores alcanzando audiencias masivas mientras que otros tienen una visibilidad considerablemente menor—, la escala logarítmica permite suavizar estas discrepancias y representar de manera más equilibrada el peso de cada actor dentro del entorno analizado.

En el caso de los medios de comunicación, la relevancia se mide a través del número de visualizaciones semanales o mensuales, extraídas mediante técnicas de *web scraping* sobre sitios web que ofrecen estos datos de manera abierta. Para los usuarios de redes sociales, la métrica utilizada es el *engagement*, que incluye interacciones como comentarios, compartidos y reacciones, proporcionando un indicador más preciso de su impacto en la difusión de narrativas.

La disposición de los actores en el mapa permite una interpretación clara de su alineación y nivel de influencia. Aquellos situados en la zona de mayor hostilidad y relevancia representan un riesgo informativo significativo, ya que tienen una gran capacidad de difusión de contenidos contrarios a la narrativa de interés. En contraste, los actores con alta relevancia y una postura neutral o favorable pueden representar aliados estratégicos o espacios de oportunidad para influir en la conversación digital.

Este tipo de visualización es esencial para la toma de decisiones estratégicas en operaciones de monitoreo informativo y gestión de la reputación digital, ya que proporciona un análisis detallado de los agentes que moldean la percepción pública y facilita el desarrollo de estrategias de comunicación adaptadas a la dinámica del ecosistema digital.

3.2 Código 2: clasificación de redes sociales y MMCC con machine learning

3.2.1 Definición de categorías y enfoque clasificatorio

El segundo módulo de la herramienta OSINT se centra en la clasificación automatizada de contenido digital en redes sociales y medios de comunicación, utilizando técnicas avanzadas de *machine learning*. Este proceso permite segmentar grandes volúmenes de datos de manera estructurada, facilitando la identificación de narrativas, tendencias y potenciales campañas de influencia dentro del entorno informativo analizado.

Dado el crecimiento exponencial del flujo de información en plataformas digitales, la capacidad de analizar y clasificar contenido de manera automatizada resulta esencial para la toma de decisiones estratégicas. La implementación de algoritmos de aprendizaje automático permite superar las limitaciones de los análisis manuales, dotando al sistema de una mayor capacidad de escalabilidad y precisión en la detección de patrones discursivos.

La eficacia del proceso de clasificación depende de una fase preliminar de análisis en la que el analista determina criterios fundamentales como:

- Selección de fuentes clave: medios de comunicación tradicionales, plataformas alternativas, redes sociales y foros especializados.
- Identificación de actores relevantes: perfiles influyentes, medios con alto impacto, comunidades digitales y usuarios anónimos con actividad significativa.
- Definición de la audiencia objetivo: usuarios afectados por la narrativa analizada y potenciales receptores de campañas informativas o de desinformación.
- Contextualización del análisis: delimitación de la zona de interés, coyuntura política, social o económica y el propósito estratégico del estudio.

Estos factores permiten establecer una taxonomía clasificatoria, asegurando que el modelo de *machine learning* se entrene con datos alineados a los objetivos del análisis. La segmentación se realiza en función de distintas categorías, que pueden organizarse bajo diversos enfoques, tales como:

- Temática del contenido: clasificación en función de áreas como política, seguridad, conflictos, economía o desinformación.
- Intención comunicativa: diferenciación entre contenido meramente informativo, artículos de opinión, discursos propagandísticos o estrategias de manipulación.
- Tipo de fuente: distinción entre medios oficiales, plataformas de comunicación alternativas, redes sociales y foros especializados.
- Relevancia y alcance: evaluación de la influencia de cada fuente dentro del entorno digital, considerando métricas como *engagement*, tráfico web y nivel de interacción.

La implementación de estos criterios permite que el modelo de *machine learning* identifique con precisión la estructura del discurso digital en la zona de interés, diferenciando entre contenido legítimo, narrativas emergentes y posibles intentos de manipulación informativa.

El gráfico circular presentado en la figura refleja la segmentación del contenido en función de las categorías predefinidas. Cada segmento representa una proporción del volumen total de publicaciones clasificadas dentro de cada tema de interés y permite al analista visualizar la distribución temática del discurso digital.

3.2.2 Algoritmos y modelos implementados

Para lograr la clasificación de las publicaciones con base en las taxonomías definidas por el analista, se ha entrenado, con una gran cantidad de datos, un modelo de *machine learning* que permita automatizar el proceso de categorización. Esto ha requerido un trabajo previo de etiquetado manual de lotes de publicaciones relativos al entorno digital objeto del análisis. Este modelo se ha ido refinando a medida que se han ido teniendo más datos para su entrenamiento y sus resultados siempre están supervisados por el analista, para determinar si la clasificación que se está haciendo de las publicaciones es correcta.

De esta forma se consigue automatizar un proceso tedioso de clasificación manual de grandes cantidades de publicaciones.

Una vez utilizado el modelo sobre un lote de publicaciones y los resultados han sido contrastados y ajustados, en el caso de que fuera necesario por el analista, se almacenan en la base de datos, con la certeza de que son completamente fiables. De esta manera, se podrá seguir refinando el modelo al entrenarlo con más lotes de datos etiquetados, de manera que, con el tiempo, la precisión de este pueda converger a un valor que se aproxime al 100 %.

3.2.3 Resultados esperados de un supuesto práctico

Una vez se clasifican las publicaciones con el apoyo del modelo de *machine learning*, la herramienta expone gráficamente los resultados para obtener una visualización de las temáticas tratadas en el entorno digital objetivo con base en las taxonomías concebidas.

En este caso, se exponen en un gráfico circular como el que se muestra en la figura siguiente. No deben sacarse conclusiones de los resultados, solo se muestran para exponer las capacidades de la herramienta.

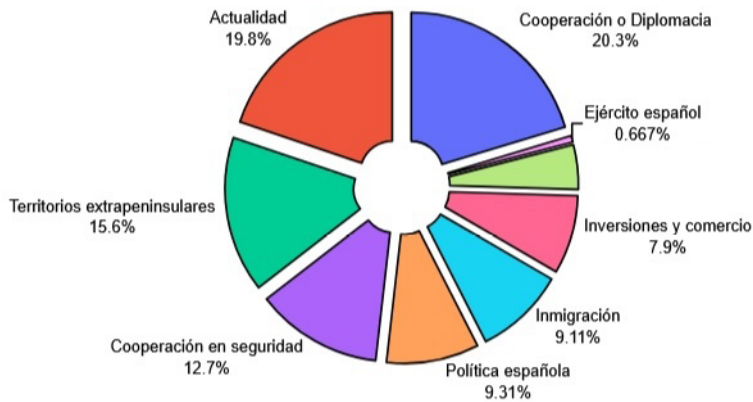


Figura 4. Gráfico de resultados de clasificación en taxonomías. Elaboración propia.

3.3 Código 3: clasificación de contenido en TikTok en varias fases

La clasificación del contenido en TikTok dentro de la herramienta OSINT desarrollada se estructura en tres fases o filtros, los cuales corresponden a distintas etapas de extracción y procesamiento de datos. Estos filtros permiten segmentar y analizar el contenido de manera progresiva, asegurando una clasificación eficiente y adaptada a los objetivos de la operación.

Antes de aplicar estos filtros, se carga en la herramienta una base de datos con usuarios previamente seleccionados. La selección de esta muestra puede variar en función del enfoque del análisis, pudiendo ser una muestra sesgada o concreta, si se quiere centrar en un nicho específico o una muestra más representativa y aleatoria, manteniendo siempre unos criterios mínimos de relevancia.

Para realizar este procesamiento de información de la red social TikTok, se han empleado métodos de *web scraping* para extraer los datos de descripción, comentarios y ubicación del vídeo de cada una de las publicaciones de los usuarios seleccionados. Además, se ha descargado y procesado del vídeo mediante un modelo de *deep learning* para la detección de objetos en imágenes.

Cada fase de clasificación de contenido desarrollada a continuación corresponde a un filtro progresivo en la detección de contenido relevante en TikTok:

3.3.1. Filtrado por *hashtags* y taxonomías en la descripción del vídeo

En la primera fase, la herramienta analiza la descripción de los vídeos publicados por los usuarios seleccionados, buscando *hashtags* o palabras clave predefinidas por el analista. Esta clasificación inicial permite una primera segmentación del contenido en función de las etiquetas que los propios creadores utilizan para categorizar sus publicaciones. El resultado de este análisis es una tabla en la que aparece el nombre de usuario, la URL de la publicación, la descripción y las palabras clave que se han encontrado en la descripción de la publicación, incluidas las que aparezcan como *hashtag*.

La información contenida en esta tabla se puede representar en gráficos como se ha hecho con los resultados de los análisis anteriores o se puede utilizar para detectar las publicaciones que sean más interesantes para analizar el entorno que esté siendo objeto de estudio. Por el momento, no se han desarrollado representaciones gráficas específicas para esta información en la herramienta que puedan ser expuestos en este artículo.

3.3.2 Filtrado por análisis de comentarios

En la segunda fase, se analizan los comentarios de los vídeos identificados en la primera fase, buscando menciones a *hashtags* o palabras clave que no estén necesariamente en la descripción, pero que puedan contextualizar el contenido. Esta fase permite ampliar el rango de detección y capturar contenido relevante que no haya sido etiquetado explícitamente por el creador del vídeo.

El resultado de este análisis consiste en una nueva columna en la tabla mencionada en el apartado anterior en la que se indica qué palabras clave y en qué cantidad se han encontrado al procesar los comentarios de la publicación.

Es importante destacar que las tres fases no son excluyentes entre sí. Si en una fase no se encuentran resultados, la herramienta sigue procesando el vídeo en los siguientes filtros hasta completar el análisis. La exclusión solo ocurre cuando, tras aplicar las tres fases, no se detecta ninguna coincidencia con los criterios establecidos.

3.3.3 Filtrado por fotogramas mediante procesado de imágenes con *deep learning*

En la tercera fase, se procede a detectar los emoticonos de interés que estén presentes en el vídeo de la publicación. Para ello, se ha entrenado un modelo de *deep learning* que detecta ciertos emoticonos en imágenes. Más concretamente, la bandera de España y de la Unión Europea. Para cada caso se deberá entrenar un modelo que detecte los emoticonos deseados.

Para poder procesar el vídeo de la publicación de TikTok, la herramienta primero obtiene la ubicación del vídeo mediante técnicas de *web scraping* y lo descarga. Después, separa el vídeo en fotogramas y los procesa uno a uno con el modelo de detección de emoticonos, escribiendo en una nueva columna de la tabla mencionada con anterioridad si ha encontrado los emoticonos de interés.

El resultado del procesamiento de un fotograma de un vídeo por el modelo *deep learning*, es de manera visual, el que se expone en la siguiente figura:

En la figura, se indica que se ha detectado un emoticono de la bandera de España con una precisión en la detección del 91,2 %. Estos datos serán los que se guarden en la fila correspondiente a la publicación junto con el análisis de palabras de la descripción y los comentarios.

Este procedimiento permite detectar mensajes inscritos en los vídeos de la red social, práctica que es bastante habitual en este entorno.

Por último, respecto a este análisis del vídeo de la publicación, hay que destacar que el entrenamiento del modelo para detectar emoticonos en los fotogramas del vídeo requiere de un trabajo previo de etiquetado de imágenes que permitan entrenar el modelo. La recopilación y el etiquetado de estas imágenes para entrenamiento son el paso más crítico para tener un modelo de detección preciso y funcional.



Figura 5. Resultado de procesamiento de detección de objetos. Elaboración propia

4 Conclusiones y aprendizajes

4.1 Impacto del programa en la organización

El desarrollo e implementación del programa OSINT basado en técnicas de PLN, inteligencia artificial y programación tradicional ha demostrado ser una herramienta muy eficiente para el análisis del entorno de la información en escenarios digitales. Su integración dentro de procesos de inteligencia ha permitido optimizar el monitoreo y clasificación de grandes volúmenes de datos, facilitando la detección de tendencias y narrativas de manera automatizada.

Uno de los principales impactos del programa ha sido la reducción del tiempo requerido para procesar grandes cantidades de información de fuentes abiertas, lo que ha mejorado la capacidad de reacción y toma de decisiones estratégicas, además de permitir un mejor aprovechamiento de los recursos humanos de la unidad. Asimismo, su modularidad y adaptabilidad han hecho posible su implementación en diversos contextos operativos, lo que lo convierte en una solución flexible y escalable.

Por otra parte, la combinación de metodologías analíticas y el conocimiento experto de analistas ha permitido afinar los resultados y mejorar la precisión del análisis, minimizando sesgos y asegurando que la información obtenida sea útil y de relevancia para la toma de decisiones.

4.2 Principales aprendizajes y contribuciones

El desarrollo del programa ha proporcionado una serie de aprendizajes clave en el ámbito de la inteligencia de fuentes abiertas y la automatización de análisis de información:

- Importancia del diseño metodológico: la eficacia del programa depende en gran medida de la correcta estructuración del proceso de recopilación, filtrado y análisis de datos. Definir criterios claros desde el inicio es esencial para garantizar la calidad de los resultados.
- Colaboración interdisciplinaria: la integración de analistas con expertos en procesamiento de datos ha sido fundamental para ajustar los modelos y mejorar su precisión. La sinergia entre conocimientos operacionales y técnicos permite desarrollar herramientas más robustas y adaptadas a necesidades específicas.
- Desafíos en la clasificación automatizada: la ambigüedad del lenguaje en redes sociales y la constante evolución de términos y expresiones requieren una actualización continua de los modelos de clasificación y análisis de sentimiento. La implementación de enfoques híbridos ha demostrado ser una solución eficaz para abordar este desafío.
- Escalabilidad y adaptabilidad: la flexibilidad del programa ha permitido su ajuste a distintos entornos y escenarios, lo que evidencia la importancia de diseñar soluciones tecnológicas modulares que puedan evolucionar junto con los requerimientos operativos.
- Impacto estratégico del análisis de datos: la capacidad de identificar patrones, tendencias y actores clave dentro del ecosistema informativo ha reforzado la utilidad del OSINT en la toma de decisiones. La detección temprana de campañas de desinformación o cambios en la percepción pública es un recurso valioso en la gestión de crisis y operaciones de seguridad.

- Combinación de distintas tecnologías: la elaboración de esta herramienta sirve como buen ejemplo de empleo de distintas tecnologías para resolver las distintas partes del problema que enfrenta. De esta forma, se ha optado por la solución más conveniente según las necesidades, coexistiendo en la herramienta tanto modelos de inteligencia artificial de *machine learning* y *deep learning* como técnicas de *web scraping* o filtrado por palabras clave realizado por algoritmos tradicionales de programación.

4.3 Próximos pasos y mejoras propuestas

Para maximizar el potencial del programa, se podrían sondear las siguientes líneas de mejora y desarrollo:

- Mejoras en la precisión del análisis de sentimiento: implementación de modelos más avanzados, para una clasificación más precisa y adaptable al contexto.
- Integración con fuentes de datos adicionales: expansión del alcance del programa mediante la incorporación de nuevas plataformas digitales y medios de comunicación, mejorando la cobertura del análisis.
- Automatización del proceso de actualización: desarrollo de un sistema que permita la actualización automática de taxonomías, términos clave y patrones de clasificación según las tendencias emergentes.
- Optimización del rendimiento y escalabilidad: mejora en la infraestructura de almacenamiento y procesamiento de datos para permitir la gestión eficiente de volúmenes crecientes de información.
- Capacitación continua del personal: fomento de programas de formación para analistas y expertos en datos, garantizando una actualización constante en técnicas de análisis y procesamiento de información.

Además, se puede tratar de combinar distintos módulos del programa con el objetivo de llegar a análisis más avanzados, como pudiera ser el uso del modelo que clasifica el sentimiento de textos con los comentarios de las publicaciones de TikTok accedidos mediante *web scraping*.

En conclusión, este programa representa un avance significativo en la aplicación de OSINT en entornos estratégicos, proporcionando una herramienta eficaz para el análisis y la toma de decisiones en contextos de crisis y seguridad. Su desarrollo y mejora continua permitirán fortalecer su impacto y ampliar su aplicabilidad en distintos ámbitos operativos. Para concluir, este desarrollo realizado con recursos limitados demuestra que las soluciones tecnológicas son accesibles y se puede sacar partido de ellas de manera autónoma si se alinean los recursos humanos disponibles de manera adecuada.