

**UNIVERSIDAD DE GUAYAQUIL**

**FACULTAD DE INGENIERIA INDUSTRIAL**

**CARRERA DE INGENIERÍA INDUSTRIAL**

**TRABAJO DE TITULACIÓN**

**PREVIO A LA OBTENCIÓN DEL TÍTULO DE**

**INGENIERO EN SISTEMAS DE INFORMACIÓN**

**ÁREA**

**CIENCIAS BÁSICAS, BIOCONOCIMIENTO Y DESARROLLO INDUSTRIAL.**

**TEMA**

**“VISUALIZACIÓN DE PATRONES DE DESINFORMACIÓN EN ECUADOR MEDIANTE UN DASHBOARD EN POWER BI”**

**AUTOR**

**BARROS SANIPATIN PAOLA LORENA**

**BARROS SANIPATIN LORENA LEONOR**

**DIRECTOR DEL TRABAJO**

**LSI. TOAPANTA BERNABE MARIUXI DEL CARMEN, MSIG.**

**GUAYAQUIL, ABRIL DEL 2025**

**Índice General**

**No Descripción Pág.**

[Introducción 1](#_Toc195001546)

Capítulo I

**Marco Teórico**

**No Descripción Pág.**

[1.1. Planteamiento del problema 2](#_Toc199861349)

[1.2. Formulación del problema 2](#_Toc199861350)

[1.3. Objeto de estudio 2](#_Toc199861351)

[1.4. Delimitación del objeto de investigación 3](#_Toc199861352)

[1.4.1. Delimitación geográfica 3](#_Toc199861353)

[1.4.2. Delimitación en Tiempo – Espacio 3](#_Toc199861354)

[1.4.3. Delimitación Semántica 3](#_Toc199861355)

[1.5. Justificación 4](#_Toc199861356)

[1.6. Alcance 5](#_Toc199861357)

[1.7. Objetivos 6](#_Toc199861358)

[1.7.1. General 6](#_Toc199861359)

[1.7.2. Objetivos específicos 6](#_Toc199861360)

[1.8. Marco Teórico 6](#_Toc199861361)

[1.8.1. Medios digitales y noticias falsas 6](#_Toc199861362)

[1.8.1.1. Impacto psicológico de las Fake News 6](#_Toc199861363)

[1.8.1.2. Desinformación en el área de la producción científica y periodística 8](#_Toc199861364)

[1.8.2. Redes sociales 9](#_Toc199861365)

[1.8.2.1. ¿A qué se dedican las dos fuentes oficiales Ecuador Chequea y Ecuador Verifica? 9](#_Toc199861366)

[1.8.2.2. Ecuador Chequea 9](#_Toc199861367)

[1.8.2.3. Ecuador Verifica 10](#_Toc199861368)

[1.8.3. Red Internacional de Verificación de Datos 12](#_Toc199861369)

[1.8.4. Corpus 18](#_Toc199861370)

[1.8.5. Google Fact Check Tools 18](#_Toc199861371)

[1.8.6. La verdad y el fact-checking 18](#_Toc199861372)

[1.8.7. La desinformación en el Ecuador 19](#_Toc199861373)

[1.8.8. Origen de fuentes de datos 20](#_Toc199861374)

[1.8.9. El poder de los datos y la visualización 21](#_Toc199861375)

[1.8.9.1. Impacto de la visualización en la toma de decisiones 21](#_Toc199861376)

[1.8.9.2. Principios de calidad de datos 21](#_Toc199861377)

[1.8.9.3. Proceso completo de los datos 22](#_Toc199861378)

[1.8.9.1. Integración de los datos 23](#_Toc199861379)

[1.9. Marco Conceptual 23](#_Toc199861380)

[1.9.1. Inteligencia de Negocios 23](#_Toc199861381)

[1.9.1.1. Componentes de la rama inteligencia de negocios 24](#_Toc199861382)

[1.9.2. Arquitectura tecnológica 28](#_Toc199861383)

[1.9.2.1. Fuentes 30](#_Toc199861384)

[1.9.2.2. Extracción, Transformación y Carga 31](#_Toc199861385)

[1.9.2.3. Almacenamiento de datos e información 31](#_Toc199861386)

[1.9.2.4. Herramientas tecnológicas para extracción de datos 34](#_Toc199861387)

[1.9.2.5. Herramientas tecnológicas para transformar datos 36](#_Toc199861388)

[1.9.2.6. Herramientas tecnológicas para base de datos 37](#_Toc199861389)

[1.9.2.7. Herramientas tecnológicas para visualizar datos 38](#_Toc199861390)

[1.10. Marco Legal 48](#_Toc199861391)

Capítulo II

**Metodología**

**No Descripción Pág.**

[2.1. Tipo de investigación 51](#_Toc199345965)

[2.1.1. Investigación descriptiva 51](#_Toc199345966)

[2.2. Enfoque de la investigación 51](#_Toc199345967)

[2.2.1. Enfoque cuantitativo 51](#_Toc199345968)

[2.2.2. Enfoque cualitativo 52](#_Toc199345969)

[2.3. Técnicas de Recolección de datos 52](#_Toc199345970)

[2.3.1. Entrevista 52](#_Toc199345971)

[2.3.2. Revisión Documental 53](#_Toc199345972)

[2.3.3. Extracción de Datos 53](#_Toc199345973)

[2.3.4. Instrumentos 54](#_Toc199345974)

[2.4. Estudio del proceso planteado 54](#_Toc199345975)

[2.4.1. Modelamiento del proceso 54](#_Toc199345976)

[2.5. Especificaciones funcionales 55](#_Toc199345977)

[2.5.1. Historia de usuarios 55](#_Toc199345978)

[2.5.2. Ficha de tareas 60](#_Toc199345979)

[2.5.3. Requerimientos funcionales 66](#_Toc199345980)

[2.5.4. Requerimientos no funcionales 67](#_Toc199345981)

[2.6. Metodología 67](#_Toc199345982)

Índice de Tablas

**No Descripción Pág.**

[**Tabla 1.** Tipos de desórdenes informativos 8](#_Toc199870577)

[**Tabla 2.** Metodología de Ecuador Chequea y Ecuador Verifica. 10](#_Toc199870578)

[**Tabla 3.** Metodología de Ecuador Chequea del 2020 11](#_Toc199870579)

[**Tabla 4.** Signatarios verificados del Código de Principios de la IFCN que se encuentran activos 13](#_Toc199870580)

[**Tabla 5.** Tipos de fuentes 20](#_Toc199870581)

[**Tabla 6.** Descripción del ciclo de la vida de los datos 22](#_Toc199870582)

[**Tabla 6.** Componentes de la rama inteligencia de negocios 24](#_Toc199870583)

[**Tabla 7.** Herramientas del Businees Intelligence 26](#_Toc199870584)

[**Tabla 8**. Diferencias Claves 27](#_Toc199870585)

[**Tabla 9**. Características de las Herramientas de Visualización 38](#_Toc199870586)

[**Tabla 10.** Tipos de herramientas para la visualización de los dashboards. 39](#_Toc199870587)

[**Tabla 11.** Arquitectura de Power BI 40](#_Toc199870588)

[**Tabla 12.** Gráficas para la visualización de los dashboards 45](#_Toc199870589)

[**Tabla 13.** Puertas de enlace 46](#_Toc199870590)

Índice de Figuras

**No Descripción Pág.**

[**Figura 1.**  Influencia de la desinformación en los periodos electorales Información tomada de(Grupo Goberna, 2025) . Elaborada por los autores. 7](#_Toc199870633)

[**Figura 2.** Calificación de Ecuador Chequea y Ecuador Verifica. Información tomada (Toapanta Bernabe et al., 2024).Elaborada por los autores. 11](#_Toc199870634)

[**Figura 3.** Comparativa de la transmisión de noticias falsas en el año 2020. Información tomada(Calero Katherine, 2021).Elaborada por los autores 20](#_Toc199870635)

[Figura 4. Ciclo de vida de los datos. Información tomada de (Ortega, 2025) 22](#_Toc199870636)

[**Figura 5.** Representación de la integración de los datos. Información Tomada de (Curto, 2020). 23](#_Toc199870637)

[**Figura 6.** Beneficios de Inteligencia Artificial. Información tomada de (Sanchez Lennin, 2024). Elaborado por los autores. 25](#_Toc199870638)

[**Figura 7.** Herramientas del BI. Tomada de (Sign Consulting, 2023). 27](#_Toc199870639)

[**Figura 8.** Modelado de datos. Tomado de (Slusarczyk Antosz, 2024). 30](#_Toc199870640)

[**Figura 9**.Arquitectura del sistema. Elaborado por los autores. 30](#_Toc199870641)

[***Figura 10.*** *Beneficios de Power Query. Información tomada de (Digari Dharmendra, 2024).Elaborada por los autores* 34](#_Toc199870642)

[**Figura 11.** Conexión a la base de datos NoSQL(MongoDB Atlas) mediante la interfaz SQL de Atlas desde Power Query Desktop. Imagen capturada por los autores. 35](#_Toc199870643)

[**Figura 12.** Elementos necesarios para incorporar MongoDB para vincular con herramientas de BI. Información tomada de (Ashraf Alameer, 2019).Elaborada por los autores. 38](#_Toc199870644)

[**Figura 13**. Arquitectura de Power Bi. Tomada de (Ferrer, 2020). 42](#_Toc199870645)

[**Figura 14.**Proceso ETL. Información tomada de (Curto, 2020). 43](#_Toc199870646)

[**Figura 15.** Configuración de la conexión. Imagen capturada desde Power BI por los autores. 44](#_Toc199870647)

[**Figura 16**. Cómo funciona la puerta de enlace. Información tomada de (Microsoft, 2025). 46](#_Toc199870648)

[**Figura 17.** Puertas de enlace. Tomada de (Microsoft Learn, 2024b). 48](#_Toc199870649)

Introducción

Actualmente, los ciudadanos están expuestos a las redes sociales y a la información que se proporciona mediante ellas. Esta información que es generada con el fin de causar miedos en ciertos casos, alteración de la realidad en que se vive y con un fin de eventos que podría perjudicar a los ciudadanos en la toma de decisiones. Por ejemplo, cuando se acerca el tiempo de elecciones electorales, la información es generada ya sea mediante la IA o es alterada para influir en el voto de los ciudadanos. Los ciudadanos que reciben la información la replican con todo el círculo que los rodea -familiares, amigos y conocidos-, realizando así una réplica que no ha sido fundamentada.

Por lo expuesto, en este proyecto se analizará los patrones de desinformación que existen en el Ecuador en las categorías de noticias como política, salud, economía, seguridad y crisis social. De acuerdo con la metodología utilizada por Ecuador Chequea y Ecuador Verifica las noticias se clasifican como: falso, engañoso, cierto, sátira, impreciso, alterado, inverificable.

En el primer capítulo se presentará la problemática general de la desinformación que existe en la plataforma X y las diversas situaciones que están expuestos los usuarios, además se abordarán la justificación, el marco teórico, el marco conceptual y el marco legal relacionado con este proyecto.

**Capítulo I**

# Marco Teórico

## Planteamiento del problema

La información maliciosa es una gran amenaza y esta práctica en Ecuador no está exenta de la problemática de la desinformación. Una gran magnitud de información engañosa o capciosa fue impulsada por diferentes plataformas digitales; así pues, Facebook, Instagram o X (Twitter) han experimentado una sensación muy significativa en la población ecuatoriana, como el resultado en la toma de decisiones informadas, la credibilidad en las instituciones políticas y el equilibrio social.

La información sobre el Ecuador está esparcida y carece de una visualización íntegra y verificable. En los sitios web de Ecuador Verifica y Ecuador Chequea-fact checkers acreditado en el país- se provee de datos útiles, para que los usuarios puedan acceder a noticias verificadas, aunque en la actualidad no se encuentran organizados de manera efectiva, para que sea más fácil analizarlos e interpretarlos.

La falta de visualización de información en el Ecuador obstaculiza el análisis de patrones temporales y geográficos, lo que restringe el acceso a la información para investigadores, corresponsales y público en general. Además, al no existir indicadores clave de desempeño impiden evaluar el impacto y la validez de la desinformación, así como también es necesario contar con herramientas que faciliten la comparación de datos históricos y equiparar las brechas de información clave en sucesos significativos.

## Formulación del problema

¿De qué manera la visualización de patrones de desinformación en Ecuador mediante un dashboard interactivo puede contribuir a la toma de decisiones informadas a los ciudadanos, periodistas y autoridades?

## Objeto de estudio

El objeto de estudio del presente trabajo de titulación es la implementación de dashboards que permita a la población ecuatoriana visualizar las tendencias de la desinformación en Google Fact Check Tools, Corpus y la red social X de Ecuador Chequea y Ecuador Verifica.

## Delimitación del objeto de investigación

### Delimitación geográfica

Este estudio se lo desarrolla en Ecuador, el cual nos ayudará a representar y examinar precisamente como se propaga la desinformación en el país por medio de la red social X de las cuentas @ECUADORCHEQUEA y @ecuadorverifica. La población ecuatoriana ha enfrentado un incremento en la desinformación en muchos casos los usuarios distribuyen contenido sin revisar la autenticidad de cada noticia lo que provoca que se propague de manera viral por lo cual es de suma importancia que la población ecuatoriana visualice estos patrones y el impacto ocasionado.

### Delimitación en Tiempo – Espacio

Este trabajo se enfocará en el análisis de patrones de desinformación implementando un dashboard interactivo en Power BI, durante una etapa de cuatro meses, desarrollado entre abril y julio del 2025. En este período se focalizará en la visualización de datos obtenidos de la plataforma X específicamente de las cuentas @ECUADORCHEQUEA y @ecuadorverifica y de Google Fact Check Tools, los datos serán almacenados y administrados en la base de datos NoSQL MongoDB.

### Delimitación Semántica

En este apartado se detallarán definiciones empleadas para el desarrollo del presente trabajo, con el propósito de simplificar la interpretación del lector. A continuación, se describe los términos importantes:

* **Visualización de datos:** Es la representación por medio de gráficas de la información y datos. Los instrumentos de la visualización de datos proveen de una forma legible de observar y captar los diferentes patrones.
* **Patrones de desinformación:** Son tendencias que se ubican en información falsa o erróneas que muchas veces son difundidas por personas que creen que son verídicas.
* **Web scraping:** Proceso automatizado para obtener información de sitios web en este caso de las publicaciones de la red social X.
* **Fact – checking:** Proceso que se basa en corroborar en artículos ya divulgados como en las redes sociales y desmentir falsedades**.**
* **Categoría:** Categoría se señala como se ha identificado el texto ya sea de la temática económica, salud pública, crisis sociales y política.
* **IFCN:** Red Internacional de Verificación de Datos que fomenta la calidad en la validación de datos. Encargada de supervisar los patrones en el campo de la comprobación de datos para aportarle al debate público y respaldar proyectos innovadores que potencien la rendición de cuentas en el periodismo.
* **Ecuador chequea:** Plataforma digital enfocada básicamente en el control de la calidad de los datos para certificar que estos sean concretos, sin ambigüedades y confiables.
* **Ecuador Verifica:** Esta plataforma tiene como objetivo es contrarrestar información engañosae impulsar la alfabetización digital.

## Justificación

Esta investigación se sostiene en la necesidad de comprender con rigor y analizar el fenómeno de la desinformación en el contexto ecuatoriano, de una forma teórica. Se apoya en marcos conceptuales de la comunicación, la desinformación y la visualización de datos y, a partir de ello, explorar cómo se da la información falsa y cómo pueden las herramientas de visualización aportar a una mejor interpretación de los fenómenos relacionados con la desinformación, además de adherirse a los estudios crecientes sobre el impacto que tiene la desinformación en las sociedades contemporáneas, aportando datos y análisis específicos del contexto ecuatoriano. En el escenario de la era digital en que cada vez más la cultura de datos cobra mayor importancia, esta investigación también se inscribe en la posibilidad de contribuir al desarrollo de la cultura de datos en el medio ecuatoriano.

La creación de un dashboard interactivo en Power BI tiene una gran importancia práctica. Este tipo de recurso será útil para medios de comunicación, investigadores y también para organizaciones de la sociedad civil, de forma que se podrá considerar como un recurso útil e importante para la toma de decisiones informadas y con las debidas consideraciones. Mediante la visualización de los patrones que permiten detectar la desinformación, el dashboard contribuirá a los intereses de las organizaciones de verificación, así podrán crear acciones para identificar y frenar la divulgación de noticias falsas. Servirá, también, para incrementar la concienciación pública sobre los peligros del fenómeno de la desinformación y la importancia que tienen la verificación de hechos. Adicionalmente, esta herramienta será de ayuda para cualquier ente gubernamental que desee lanzar campañas para la prevención de la desinformación.

Desde el punto de vista metodológico, el presente trabajo ejemplifica el uso de Power BI como herramienta para la visualización y análisis de datos complejos en el campo de la comunicación; se da lugar a una forma rigurosa de almacenar, estructurar y analizar los datos, asegurando la validez y la fiabilidad de los resultados. Al mismo tiempo que se explorarán y aplicarán técnicas de visualización de datos interactivos que facilitarán la exploración y comprensión de los patrones de la desinformación. También se espera que la metodología y los resultados sean reproducibles para que otras personas investigadoras puedan utilizar y ampliar el trabajo en el estudio de la desinformación. La utilización de fuentes oficiales como Ecuador Chequea, una iniciativa dedicada a la verificación de hechos y al combate de la desinformación en el ámbito público, y Ecuador Verifica, una plataforma que se especializa en la detección y análisis de noticias falsas y contenido engañoso en Ecuador, corroboran la credibilidad y la validación del trabajo de investigación.

## Alcance

El alcance de este proyecto es descriptivo, se limita a la recopilación y análisis de datos de desinformación provenientes exclusivamente de las siguientes fuentes Ecuador Chequea y Ecuador Verifica, Google Fact Check Tools y corpus de noticias verificadas, utilizando Power BI para desarrollar un dashboard interactivo. En él mismo, se van a visualizar patrones temporales y geográficos que permita la segmentación de noticias por tema, fuente, fecha y tipo de calificación, muestre indicadores clave de desempeño (KPIs) sobre la cantidad de noticias verificadas y tendencias, y ofrezca la funcionalidad de comparación histórica, todo dentro de las capacidades estándar de Power BI y los datos estructurados disponibles, enfocándose en temas consistentemente abordados por las fuentes y en un período de tiempo definido desde Enero 2020 y Julio 2025, para un público objetivo de investigadores, periodistas y el público en general, sin incluir análisis de desinformación en redes sociales no verificada por las fuentes oficiales ni el desarrollo de algoritmos de detección automática.

Este rango de tiempo permitirá analizar la evolución de la desinformación en un periodo significativo, identificando tendencias y patrones a lo largo del tiempo. El dashboard incluirá funcionalidades que permitirán visualizar patrones temporales, segmentar noticias por tema, fuente, fecha y tipo de calificación, y presentar indicadores clave de desempeño (KPIs) relacionados con la cantidad de noticias verificadas y las tendencias de la desinformación. Además, se incluirá una funcionalidad de comparación histórica para facilitar el análisis de la evolución de la desinformación.

## Objetivos

### General

Desarrollar un dashboard interactivo en Power BI que permita el análisis y visualización de patrones de desinformación en Ecuador, proporcionando información clave sobre la evolución y distribución de noticias verificadas en el país.

### Objetivos específicos

1. Implementar un proceso de extracción, limpieza y transformación de datos para garantizar la calidad y consistencia de la información.
2. Implementar filtros y segmentaciones que permitan a los usuarios profundizar en los datos y descubrir patrones específicos.
3. Desarrollar visualizaciones claras y concisas para mostrar los KPIs, permitiendo a los usuarios comprender rápidamente la situación de la desinformación en el país.
4. Publicar el Dashboard en Power BI para facilitar el acceso y la interpretación de datos sobre desinformación en Ecuador, promoviendo la toma de decisiones informadas entre los usuarios

## Marco Teórico

### Medios digitales y noticias falsas

Para (Pineda Solorio, 2024) las “fake news” son conocidas como información que ha presentado de manera manipulada para distorsionar o engañar, su difusión se debe a que los algoritmos de las redes sociales otorgan mayor importancia al contenido viral y sensacionalista, sin poseer un criterio claro entre la autenticidad y la falsedad. Diversos estudios han evidenciado que las noticias falsas se distribuyen con mayor premura que las verdaderas en plataformas como X, ya que originan mayor influencia emocional.

“Fake news se refiere a la información deliberadamente falsa o engañosa, creada con la intención de desinformar o manipular a la audiencia”.(Pineda Solorio, 2024)

Según lo referenciado por Castillo-Riquelme et al. (2021), la información engañosa alude a la elaboración o distorsión intencionada de un dato comprobable para defraudar al público.

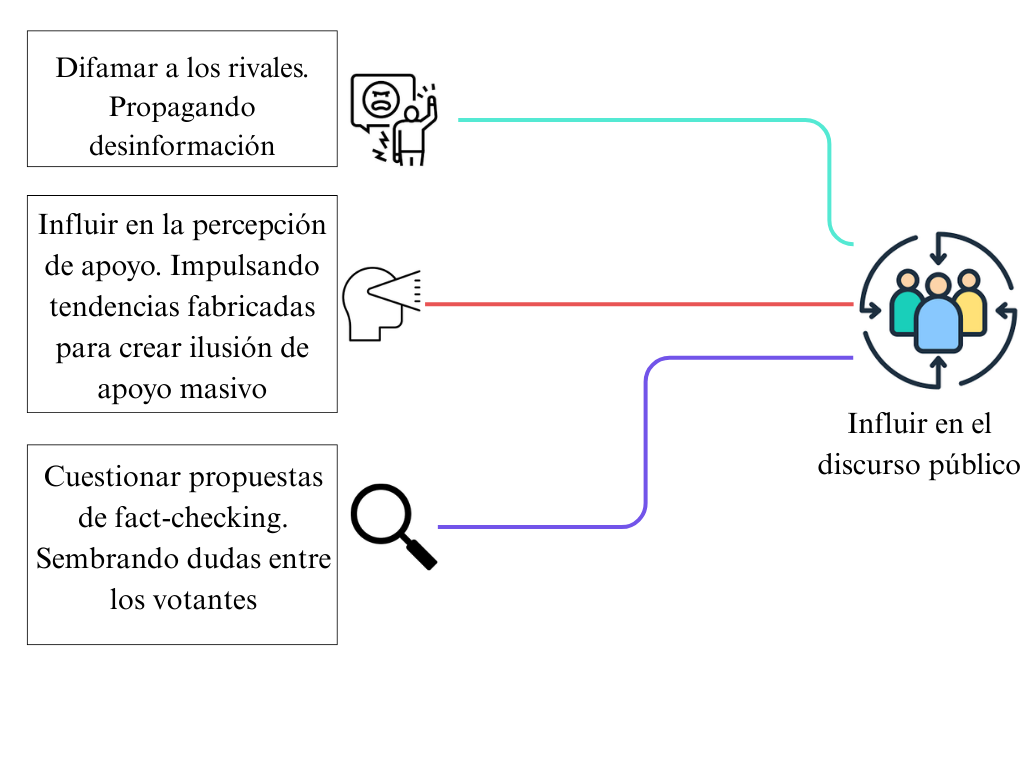
#### Impacto psicológico de las Fake News

Según (Pineda Solorio, 2024)

Las fake news tienen un impacto psicológico significativo en los usuarios. De inicio es el "efecto de verdad ilusoria" desencadena que la divulgación de la información, incluso si es falsa, intensifique la posibilidad de que se conceda como fidedigna. Esta influencia se potencia en las redes sociales por ejemplo X (antes llamada Twitter), donde la información se replica de manera continua.

Además, el sesgo de verificación es un aspecto fundamental debido a que los individuos exploran y aceptan la información que intensifica sus convicciones preexistentes. La información tergiversada, por lo tanto, explotan este sesgo con el fin de involucrarse en las emociones y en los principios de la audiencia, lo que impide que los usuarios replanteen la autenticidad de aquello que adquieren en las redes sociales. En efecto, esto no solo refuerza posturas imprecisas, sino que también provocan descontento en relación con los medios de comunicación y las instituciones legítimas.

Al estar vulnerable frecuentemente a noticias falsas, se tiene como efecto que estaría en la posibilidad de desnaturalizar la percepción de la realidad, perjudicando así las múltiples actividades como procesos electorales y la participación cívica. El efecto de verdad ilusoria como el sesgo de confirmación se integran para afianzar ideas erróneas, haciendo que resulte más complejo rectificar la desinformación. (Pineda Solorio, 2024)



**Figura 1.**  Influencia de la desinformación en los periodos electorales Información tomada de(Grupo Goberna, 2025) . Elaborada por los autores.

#### Desinformación en el área de la producción científica y periodística

El rol de los medios de comunicación se ha cuestionado tanto por parte de los diferentes actores políticos y sociales debido a su histórica relación con grupos políticos, empresariales y sectores con influencia en el devenir de la sociedad. Por esta razón, la supuesta objetividad que buscan transmitir en su rol de gestores puesta en duda.

**Conceptualización de la Desinformación**

**Tabla 1.** Tipos de desórdenes informativos

|  |  |
| --- | --- |
| Tipo de Desinformación | Explicación Sencilla |
| Sátira y Parodia | Algo que intenta ser gracioso, pero que a veces puede confundir a la gente y hacerles creer cosas que no son verdad. |
| Contenido Engañoso | Cuando se usa la información de manera tramposa para que la gente piense de cierta forma sobre algo. |
| Contenido Impostor | Cuando alguien se hace pasar por una fuente real de información que es de confianza. |
| Contenido Fabricado | Inventar noticias o datos completamente falsos para engañar a la gente y hacer daño. |
| Conexión Falsa | Cuando el título de una noticia, las fotos o los textos que acompañan a una imagen no tienen nada que ver con la información que se está dando. |
| Contexto Falso | Compartir información que es verdadera, pero añadiéndole detalles o una historia que no es cierta. |
| Contenido Manipulado | Cambiar o retocar fotos o información real para que parezca otra cosa |

*Información tomada del* (Sánchez Duarte & Magallón-Rosa, 2023)*, Elaborado por los autores.*

### Redes sociales

Herramienta que conecta a una amplia cantidad de usuarios en todo el mundo para mantenerlos comunicados de manera online (César Bartolomé, 2021).

Entre 2004 y 2010, se han visto surgir cinco de los servicios más reconocidos, utilizados y extendidos en el ámbito digital como Facebook, YouTube, WhatsApp, Instagram y Twitter. Estas plataformas se han visto como cada día han ido aumentado y se han convertido en líderes en su dominio y llegando a afectar masivamente la interacción pública y el acceso a la información en línea (César Bartolomé, 2021).

#### ¿A qué se dedican las dos fuentes oficiales Ecuador Chequea y Ecuador Verifica?

Conforme a lo indicado en la página de Ecuador Chequea (2022), se dedican a hacer periodismo, pero con un enfoque muy fuerte en verificar la información. Priorizan la verificación de datos que circulan en redes sociales, porque es donde más se difunde la desinformación. Además, revisan lo que expresan los funcionarios y políticos, para garantizar que la información sea correcta.

Un aspecto relevante que mencionan es que no verifican opiniones ni promesas, solo datos que se pueden comprobar. Ejercen el rigor periodístico y tienen sus propias directrices editoriales, que se puede consultar en su página web y en redes sociales.

De acuerdo con la información revisada en Ecuador Verifica (2020) emplea herramientas de verificación de datos, lo que conocemos como “fact-checking”. Básicamente, se dedican a comprobar si lo que dicen los candidatos es verdad, mentira o necesita más contexto. A su vez se concentran en verificar la información que se difunde en redes sociales y mensajería instantánea, especialmente en el periodo electoral, con el fin de evaluar su precisión.

Un aspecto clave es que, para llevar a cabo estas verificaciones, se apoyan en la metodología de Ecuador Chequea.

#### Ecuador Chequea

La función de chequeado de información indaga en el discurso público con hechos y datos disponibles para así garantizar la veracidad a los ciudadanos ecuatorianos. Ante la avalancha de información, donde es difícil distinguir lo verdadero de lo falso, estas plataformas son cada vez más necesarias. Sirven como un espacio para noticias comprobadas y relatos profundos, creados con principios éticos y pensando en el bienestar de la gente (Palacios y Cusot, 2019)

Desde enero de 2019 Ecuador chequea forma parte de la International Fact Checking Network y suscribe sus principios (Ecuador Chequea, 2022).

#### Ecuador Verifica

Ecuador Verifica surgió después de Ecuador Chequea, un sitio web ecuatoriano que ya desde 2006 se dedicaba a combatir la desinformación. Ambos pertenecen a la misma organización, Fundamedios.

Ecuador Verifica, aunque lleva años funcionando, no es miembro de la IFCN(The International Fact-Checking Network). Para ser parte de esta organización, se deben cumplir varios requisitos, incluyendo publicar una verificación por semana durante seis meses antes de solicitar la membresía (G. Vélez Bermello, 2024).

**Tabla 2.** Metodología de Ecuador Chequea y Ecuador Verifica.

|  |  |
| --- | --- |
| Calificación | Descripción |
| Cierto | El contenido que es cierto es aquel que muestra información que sea precisa y que además tenga congruencia con los datos donde se ha obtenido la información deben de ser fuentes verificadas. Lo cual significa que, las declaraciones se presentan con exactitud y no se eliminan ninguna parte, ni se limitan la contextualización. |
| Falso | El contenido no tiene una fuente que presente la información de manera verídica y confiable. Esto se da cuando declaración es incoherente con la información que se proporciona. |
| Alterado | Es aquel que dicho contenido que se presenta ya sea en medios audiovisuales alterado que podría causar engaño a las personas. Esta categoría lo que integrar son fotografías alteradas, audios que no es verificada y que es catalogada como falsa, videos, cadenas de WhatsApp replicadas. |
| Engañoso | Este contenido que presenta no se lo considera en su totalidad como falso. Es engañoso cuando la afirmación que se hace tiene congruencia con ciertos aspectos de la información presentada al público o que presenta un poco de coincidencias una parte de ella con ciertos datos, pero a su vez demuestra que ha sido alterada con la finalidad de engañar al público en general. |
| Impreciso | Es impreciso cuando la aseveración es consistente con los datos disponibles, pero esta información se omite u oculta ciertas partes de la publicación. |
| Sátira | Este contenido es aquel que muestra la información de tres formas que son de manera de exageración, ironía, ridiculización. Además, este es más usado en temas como la política, en temas de la religión o temas sociales. |
| Inverificable | Es aquella afirmación donde no se ha podido identificar las fuentes oficiales o argumentos que fundamenten dicho pronunciamiento. |

*Información tomada de (Ecuador Chequea, 2022; Ecuador Verifica, 2020), Elaborado por los autores.*



**Figura 2.** Calificación de Ecuador Chequea y Ecuador Verifica. Información tomada (Toapanta Bernabe et al., 2024).Elaborada por los autores.

En el 2020 existían otro tipo de categorización de la metodología en Ecuador Chequea a continuación presentaremos esta tabla con sus respectivos conceptos:

**Tabla 3.** Metodología de Ecuador Chequea del 2020

|  |  |
| --- | --- |
| Categoría | Concepto |
| Cierto | Manifestaciones de sujetos que son precisas en todos sus aspectos, alineadas con fuentes fidedignas, sin omisión de datos ni restricción en la contextualización. |
| Sí, pero | Declaraciones con información fragmentada con capacidad de verificación, que solicita mayor interpretación contextual para establecer su rango de acierto. |
| Insostenible | Enunciados sin fuente clara y sin sesgos. No está en condiciones de evaluar su fiabilidad sin una fuente certera. |
| Falso | Informe que discrepa explícitamente de los datos objetivos. Es una falacia certificada. |
| Falseta | Contenidos difundidos masivamente que corresponden a desinformación, como imágenes corrompidas, audios modificados, videos descontextualizados, cadenas de WhatsApp, etc. |
| A profundidad | Nueva categoría integrada tras el comienzo de la pandemia. Empleada para separar las noticias seguidamente de una evaluación minuciosa. |

*Información tomada de (G. L. Vélez Bermello, 2020)*

### Red Internacional de Verificación de Datos

Según la página de Poynter (2025):

La Red Internacional de Verificación de Datos (IFCN) nació en el año del 2015 y su propósito es articular la comunidad en crecimiento de los verificadores de datos a nivel mundial, con ello contribuyendo así a la integridad de la información en la lucha contra la desinformación que se lleva a cabo a nivel mundial y apoyando y ayudando a los verificadores mediante la creación de redes, el fortalecimiento de capacidades y el trabajo en colaboración.

Además, indica que la IFCN está relacionada con más de 170 organizaciones y con ellas se llevan a cabo la verificación de datos alrededor de todo el mundo. El grupo de trabajo que ellos tienen se mantiene alerta y en contacto con las tendencias del campo de la verificación de datos para ofrecer recursos a los verificadores y participar de los debates públicos y apoyar nuevas ideas que ayuden a promover la responsabilidad en el periodismo.

**Tabla 4.** Signatarios verificados del Código de Principios de la IFCN que se encuentran activos

|  |  |  |
| --- | --- | --- |
| Organización | País | Verificado en |
| 20 Minutes Fake off | Francia | 23/8/2023 |
| ABS-CBN Corporation | Filipinas | 27/3/2024 |
| AFP fact checking | Francia | 25/9/2024 |
| AP Fact Check | Estados Unidos | 23/8/2023 |
| APA - Austria Presse Agentur | Austria | 13/3/2024 |
| Africa Check | Sudáfrica | 8/10/2024 |
| Animal Político - El Sabueso | México | 23/8/2024 |
| Aos Fatos | Brasil | 13/3/2024 |
| Australian Associated Press | Australia | 12/3/2024 |
| BOOM | India | 19/1/2024 |
| Beam Reports | Sudán | 27/2/2025 |
| Belarusian Investigative Center | República Checa | 18/6/2024 |
| CORRECTIV | Alemania | 19/1/2024 |
| CableCheck | Nigeria | 18/3/2025 |
| Cazadores de Fake News | Venezuela, República Bolivariana de | 14/3/2025 |
| Cek Fakta - Liputan 6 | Indonesia | 19/1/2024 |
| Check Your Fact | Estados Unidos | 13/3/2024 |
| Chequeado | Argentina | 7/10/2024 |
| CivilNet | Armenia | 4/10/2024 |
| Colombiacheck | Colombia | 27/3/2024 |
| Cotejo.Info | Venezuela, República Bolivariana de | 4/3/2024 |
| Demagog Association | Polonia | 4/10/2024 |
| Demagog.cz | República Checa | 22/11/2024 |
| Demagog.sk | Eslovaquia | 26/3/2025 |
| Deutsche Welle | Alemania | 19/3/2024 |
| Digital Forensics, Research and Analytics Centre (D-FRAC) | India | 27/3/2024 |
| Digiteye | India | 8/10/2024 |
| Doble Check (Radioemisoras UCR) | Costa Rica | 18/3/2025 |
| Doğrula | Turquía | 31/1/2024 |
| Doğruluk Payı | Turquía | 7/12/2023 |
| EFE Verifica | España | 10/5/2024 |
| Ecuador Chequea | Ecuador | 5/12/2024 |
| Efecto Cocuyo (Cocuyo Chequea) | Venezuela, República Bolivariana de | 26/3/2025 |
| Ellinika Hoaxes (Greek Hoaxes) | Grecia | 13/3/2024 |
| Encosta no Horizonte, Lda | Portugal | 26/3/2025 |
| FACTLY MEDIA & RESEARCH | India | 10/5/2024 |
| Fact Check Cyprus | Chipre | 23/8/2024 |
| Fact Check Zimbabwe | Zimbabue | 27/2/2025 |
| Fact Investigation Platform (FIP.am) | Armenia | 18/3/2025 |
| FactCheck Georgia | Georgia | 19/3/2024 |
| FactCheck.org | Estados Unidos | 13/3/2024 |
| FactCheckNI | Reino Unido | 27/2/2025 |
| FactCrescendo | India | 13/5/2024 |
| FactReview | Grecia | 18/6/2024 |
| FactWatch | Bangladesh | 18/6/2024 |
| Facta | Italia | 27/2/2025 |
| Factcheck Lab | Hong Kong | 18/6/2024 |
| Factcheck.bg | Bulgaria | 23/8/2024 |
| Factcheck.kz | Kazajistán | 11/11/2024 |
| Factchequeado.com | Estados Unidos | 27/2/2025 |
| FakeNews Tragač | Serbia | 25/9/2024 |
| Fakt Yoxla | Azerbaiyán | 27/2/2025 |
| Faktisk.no | Noruega | 2/11/2023 |
| Faktograf-udruga za informiranu javnost | Croacia | 4/6/2024 |
| Faktoje.al | Albania | 22/11/2024 |
| Fast Check CL | Chile | 13/2/2024 |
| Ferret Fact Service | Reino Unido | 31/1/2024 |
| First Check | India | 10/5/2024 |
| France 24 - Les Observateurs | Francia | 5/12/2024 |
| Full Fact | Reino Unido | 19/1/2024 |
| Fundacja "Przeciwdziałamy Dezinformacji" | Polonia | 10/5/2024 |
| Funky Citizens | Rumania | 22/11/2024 |
| GWARA MEDIA | Ucrania | 5/12/2024 |
| Greece Fact Check | Grecia | 22/1/2024 |
| INTERNEWS KOSOVA | Kosovo | 16/10/2023 |
| India Today Fact Check | India | 8/10/2024 |
| Istinomer | Serbia | 13/3/2024 |
| Istinomjer | Bosnia y Herzegovina | 13/3/2024 |
| Knack Magazine, Roularta Media Group | Bélgica | 10/5/2024 |
| La Silla Vacía | Colombia | 4/10/2024 |
| Lakmusz.hu - Magyar Jeti Zrt. | Hungría | 25/9/2024 |
| Lead Stories | Estados Unidos | 7/12/2023 |
| Les Surligneurs | Francia | 23/8/2024 |
| Local Voices Media Network | Liberia | 10/12/2024 |
| Lupa | Brasil | 19/1/2024 |
| MAFINDO | Indonesia | 27/3/2024 |
| Mala Espina Check | Chile | 10/5/2024 |
| Maldita.es | España | 2/11/2023 |
| MediaWise | Estados Unidos | 27/2/2025 |
| Medical Dialogues | India | 14/3/2025 |
| Medizin transparent - Universität für Weiterbildung Krems (Donau-Universität Krems) | Austria | 31/1/2024 |
| Metamorphosis Foundation | Macedonia del Norte | 22/11/2024 |
| MindaNews | Filipinas | 26/8/2024 |
| MyGoPen | Taiwán | 5/12/2024 |
| Myth Detector | Georgia | 10/5/2024 |
| NepalFactCheck.org | Nepal | 27/2/2025 |
| Nest center for Journalism Innovation and Development NGO | Mongolia | 23/8/2024 |
| NewsMobile | India | 27/2/2025 |
| Newschecker | India | 8/10/2024 |
| Newsmeter (Fifth Estate Digital Private Limited) | India | 18/6/2024 |
| Newtral | España | 25/9/2024 |
| Observador - Fact Check | Portugal | 13/3/2024 |
| Ocote | Guatemala | 8/3/2024 |
| Open.online | Italia | 27/2/2025 |
| Oštro, center for investigative journalism in the Adriatic region | Eslovenia | 19/1/2024 |
| PA Media | Reino Unido | 27/2/2025 |
| PT. Kompas Cyber Media | Indonesia | 10/12/2024 |
| Pagella Politica | Italia | 12/9/2023 |
| Palestinian platform for Fact-checking and Media Literacy "Kashif" | Territorio Palestino, Ocupado | 10/12/2024 |
| Patikrinta 15min | Lituania | 22/11/2024 |
| PesaCheck | Kenia | 18/6/2024 |
| Piga Firimbi | Kenia | 10/12/2024 |
| PolitiFact | Estados Unidos | 19/1/2024 |
| Polígrafo | Portugal | 5/12/2024 |
| Pravda | Polonia | 27/2/2025 |
| Premium Times Center for Investigative Journalism | Nigeria | 13/3/2024 |
| PressOne.PH | Filipinas | 10/5/2024 |
| Provereno.Media | Estonia | 5/12/2024 |
| RMIT Lookout | Australia | 17/12/2024 |
| Rappler | Filipinas | 3/11/2023 |
| Raskrikavanje | Serbia | 27/2/2025 |
| Raskrinkavanje | Bosnia y Herzegovina | 27/2/2025 |
| Raskrinkavanje.me | Montenegro | 18/6/2024 |
| Reuters | Estados Unidos | 27/1/2025 |
| Science Feedback | Francia | 10/5/2024 |
| Snopes.com | Estados Unidos | 19/1/2024 |
| StopFake.org | Ucrania | 18/6/2024 |
| Stopfals.md | Moldova, República de | 31/1/2024 |
| Suara.com | Indonesia | 10/9/2024 |
| T Verifica — Noticias Telemundo | Estados Unidos | 13/3/2024 |
| THIP Healthtech Pvt Ltd | India | 8/10/2024 |
| TOGOCHECK | Togo | 27/8/2024 |
| TV Today Network Limited - The Lallantop | India | 13/3/2024 |
| Taiwan FactCheck Center | Taiwán | 30/8/2023 |
| Telugupost.com | India | 8/10/2024 |
| Tempo.co | Indonesia | 10/5/2024 |
| Teyit | Turquía | 23/8/2024 |
| The Canadian Press | Canadá | 10/5/2024 |
| The Dispatch | Estados Unidos | 13/3/2024 |
| The Journal FactCheck | Irlanda | 4/10/2024 |
| The Quint | India | 5/12/2024 |
| The Stage Media -Liberia | Liberia | 23/8/2024 |
| The Washington Post Fact Checker | Estados Unidos | 23/8/2024 |
| The Whistle | Israel | 23/1/2025 |
| Tirto ID | Indonesia | 31/1/2024 |
| TjekDet.dk | Dinamarca | 23/8/2024 |
| UOL Confere | Brasil | 27/2/2025 |
| USA TODAY | Estados Unidos | 18/6/2024 |
| VRT NWS | Bélgica | 27/2/2025 |
| Verafiles Incorporated | Filipinas | 7/12/2023 |
| Verificador de La República | Perú | 23/8/2024 |
| Verificat | España | 18/6/2024 |
| Verify Media Platform | Turquía | 5/12/2024 |
| Viral Check | Portugal | 4/10/2024 |
| Vishvas News | India | 23/8/2024 |
| VoxUkraine | Ucrania | 13/3/2024 |
| Wisconsin Watch | Estados Unidos | 22/11/2024 |
| Wojownicy Klawiatury | Polonia | 31/1/2024 |
| Youturn | India | 18/6/2024 |
| Dpa Deutsche Presse-Agentur GmbH | Alemania | 19/1/2024 |
| franceinfo.fr | Francia | 10/5/2024 |

*Información tomada de (International Fact-Checking Network, 2025). Elaborado por los autores*.

### Corpus

Según en el artículo de (José M. García-Miguel, 2022)

La lingüística de corpus se basa en un conjunto de métodos que involucran la recopilación y análisis de corpus en estudios lingüísticos tanto teóricos como prácticos. Lo que distingue a esta rama es el uso de corpus textuales como principal fuente de datos. Un corpus se puede describir como una recopilación de textos orales o escritos creados en un contexto comunicativo real, que representan una lengua o variante lingüística, almacenados digitalmente para su posterior análisis lingüístico.

### Google Fact Check Tools

Según la página oficial (Google, n.d.) las herramientas destinadas a la verificación de datos incluyen dos componentes principales: el Explorador de Verificaciones de Datos y la Herramienta de Marcado de Verificaciones de Datos. Estas herramientas están diseñadas para facilitar el trabajo de periodistas, investigadores y verificadores en sus labores. Cabe destacar que Google no participa en la creación ni respalda directamente estas verificaciones.

### La verdad y el fact-checking

El periodismo tradicionalmente revisa la información antes de publicarla. Por otro lado, el “fact-checking” se dedica a comprobar lo que dicen los políticos y los datos oficiales, sin importar si eso salió en un medio o en redes sociales. Básicamente, usa técnicas del periodismo de datos para mostrar cuando hay errores, cosas confusas, mentiras o información incorrecta (Vélez Bermello y Henríquez Coronel, 2024).

El “fact-checking” toma información que ya ha sido publicada y la revisa a fondo. Se comparan datos, cifras y fuentes para determinar si la información es correcta. Los resultados se clasifican usando etiquetas como “verdadero”, “falso”, “dudoso” o “insostenible”, dependiendo de la plataforma (Vélez Bermello y Henríquez Coronel, 2024)

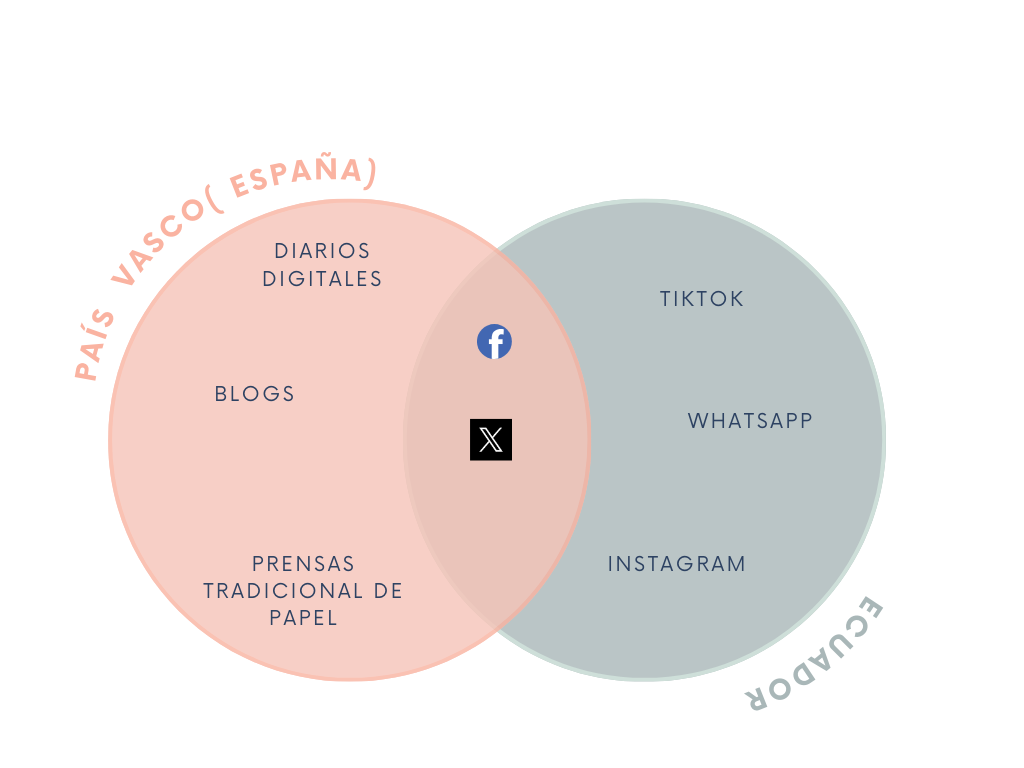
### La desinformación en el Ecuador

La desinformación se ha transformado en un evento significativo en la era digital, perjudicando la opinión pública y repercutiendo en la toma de decisiones en diferentes áreas, como por ejemplo en la política y la salud. En el Ecuador, el tránsito de la información engañosa ha causado intranquilidad en torno a la conmoción de dicha desinformación en la democracia y en la cohesión social (Ulchur Iván, 2024). El propósito de este marco teórico es proporcionar una referencia conceptual y empírica que sustente la investigación en torno a la visualización de patrones de desinformación en Ecuador, a través de un tablero interactivo en Power BI.

La desinformación hace referencia a la tergiversación de datos falsos o engañosos, ya sea de forma deliberada o no voluntaria. Según el Reporte Digital sobre las noticias 2022 de Reuters, la circulación de noticias falsas sigue persistiendo en una cuestión significativa en América Latina, donde el 60% de los encuestados sostiene haber localizado información incierta en internet (Newman et al., 2022).Asimismo, este incidente ha percibido gran interés en el auge de las redes sociales, en el cual la información se extiende aceleradamente.

En el Ecuador, la difusión de información falsa ha sido objeto de estudio en diversos trabajos de investigación. Según un informe de la Defensoría del Pueblo, elaborado por Tusa y Durán (2019), el país ha vivido un aumento de la circulación de noticias falsas, y, sobre todo, en periodos electorales.

Al presentar los datos de forma accesible se podría propiciar una mejora en la alfabetización mediática y con ello se podría empoderar a los ciudadanos para que puedan tomar decisiones bien fundamentadas.



**Figura 3.** Comparativa de la transmisión de noticias falsas en el año 2020. Información tomada(Calero Katherine, 2021).Elaborada por los autores

### Origen de fuentes de datos

Los datos pueden ser internos o externos. Los datos que son internos son aquellos que proviene de la propia empresa, y pueden presentarse de diferentes maneras ya sea por medio de bases de datos, archivos o reportes de sistemas informáticos empresariales como ERP, CRM, SCM (Slusarczyk Antosz, 2024).

Los datos que son externos son aquellos que provienen ya sea por medio de redes sociales, sistemas de información geográfica, sistemas gubernamentales o por medio de web services, entre otros (Slusarczyk Antosz, 2024).

A continuación, se presentan varios tipos de fuentes:

**Tabla 5.** Tipos de fuentes

|  |  |  |
| --- | --- | --- |
| Categoría | Descripción | Ejemplos |
| Archivos de datos | Archivos que contienen datos en diferentes formatos. | Texto, Excel, CSV, JSON, XML. |
| Bases de datos | Almacenan datos estructurados y no estructurados, tanto locales como en la nube. | Access, SQL, Oracle, Azure SQL DB, NoSQL. |
| Servicios de almacenamiento en línea | Son aquellas plataformas que están dedicadas a almacenar y acceder a datos desde Internet. | Google Drive, Dropbox, OneDrive, Box, Zoho Docs. |
| Aplicaciones empresariales | Son herramientas que se las utilizan para llevar a cabo gestiones y analizar datos en entornos empresariales. | Sistemas bancarios, POS, ERP, CRM, Google Analytics. |
| Recursos de Internet | Datos que son recuperados directamente de sitios web o servicios en línea. | Sitios web, fuentes Web, servicios en línea. |
| Aplicaciones personalizadas | Conectores de datos personalizados mediante API para integrar datos de diversas aplicaciones. | DataChannel, Cervinodata, Salesforce, SYNCHUB. |

*Información tomada de (Slusarczyk Antosz, 2024). Elaborado por los autores*.

### El poder de los datos y la visualización

El impacto del big data radica en el procesamiento y evaluación de datos masivos de información. A fin de que estos sean provechosos, es fundamental depurarlos, interpretarlos y analizarlos(Esther Paniagua, 2021). En esta etapa la visualización de datos, como la que ofrece Power BI desempeña un papel fundamental ya que influye en el proceso del big data permitiendo a los usuarios sin formación técnica a comprender patrones, tendencias y obtener información valiosa que de otro modo resultarían difíciles de manejar grandes volúmenes de información.

Al hacer los datos asequibles y claros, Power BI permite tomar “decisiones respaldadas por datos”.

#### Impacto de la visualización en la toma de decisiones

La visualización de datos interactiva contribuye a que la toma de decisiones sea más ágil y óptima porque facilita a los usuarios acceder en tiempo real a datos relevantes y les brinda la oportunidad de evaluar varios escenarios y opciones (Zhang, 2024).

La visualización de datos interactiva minimiza la carga innecesaria cognitiva asociada con la interpretación de datos complejos, potenciando la capacidad analítica de quienes toman decisiones para implementar estimaciones precisas y actuar con rapidez (Zhang, 2024).

#### Principios de calidad de datos

La calidad de datos se entiende como el nivel en el que los datos satisfacen con los requisitos para ser empleados de forma óptima. Los principios fundamentales incluyen la precisión, completitud, consistencia, relevancia y puntualidad. Los datos de calidad superior son indispensables para decidir con conocimiento de causa y prevenir errores(Sarango et al., 2025).

#### Proceso completo de los datos

Se alude a las diversas fases por las que transcurren los datos desde su desarrollo hasta su depuración. Estas fases son necesarias para controlar los datos de manera productiva y asegurar su nivel y protección a lo largo del periodo.

Imagen que contiene Diagrama

El contenido generado por IA puede ser incorrecto.

Figura 4. Ciclo de vida de los datos. Información tomada de (Ortega, 2025)

**Tabla 6.** Descripción del ciclo de la vida de los datos

|  |  |
| --- | --- |
| Etapa | Descripción |
| Ingesta y almacenamiento | Los datos tienen la posibilidad de estar esparcidos en distintos bases de datos, repositorios, sistemas.  Su función es agrupar los datos de forma  recurrente y consolidarlos en un solo repositorio o base de datos. |
| Procesamiento y preparación | En esta etapa se suele depurar registros duplicados o inválidos  La finalidad es adquirir datos lo suficientemente congruentes para lograr examinarlos y efectuar diferentes tipos de análisis |
| Exploración | Se descomponen los datos para identificar patrones, tendencias y relaciones |
| Experimento y predicción | Con los datos accesibles, se ejecutan experimentos a modo de entrenamiento del modelo para visualizar cual es el mejor |

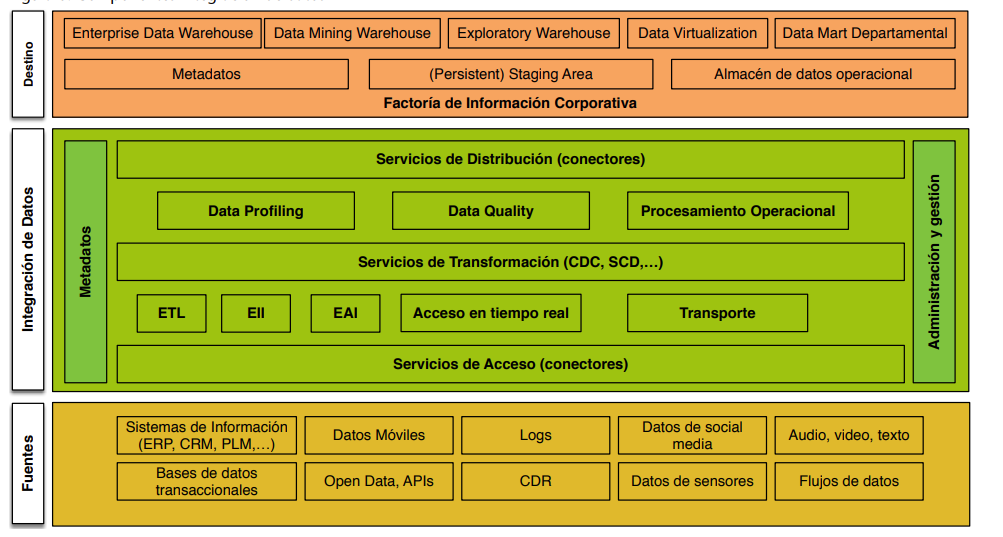
*Información tomada de (Ortega, 2025). Elaborada de los autores.*

#### Integración de los datos

Se comprende por integración de datos al conjunto de herramientas que se pueden utilizar para así obtener una perspectiva única y consistencia de los datos del negocio. Estas herramientas son aplicaciones, productos, técnicas y tecnologías.(Curto, 2020)

La integridad datos tiene múltiplos propósitos:

* Sincronizar bases de datos, esto sirve para que la información este actualizada y sea consistente con todos los sistemas.
* Otro propósito es consolidad sistemas, esto hace referencia a unir datos de diversas plataformas para así crear una visión única y centralizada
* Como ultimo propósito se tiene proveer datos para proyectos de BI.



**Figura 5.** Representación de la integración de los datos. Información Tomada de (Curto, 2020).

## Marco Conceptual

### Inteligencia de Negocios

En la actualidad según María Slusarczyk (2024) se entiende por inteligencia de negocios aquello que permite emplear estrategias y herramientas para convertir datos sin procesar en información útil, e información en conocimiento valioso, con el fin de optimizar el proceso de toma de decisiones en una organización. Además, con esto se busca que las decisiones estén fundamentadas en criterios racionales. Es esencial maximizar el aprovechamiento que las empresas hacen de sus datos y la información que poseen.

A continuación, se menciona 4 tipos de fuentes primordiales y herramientas:

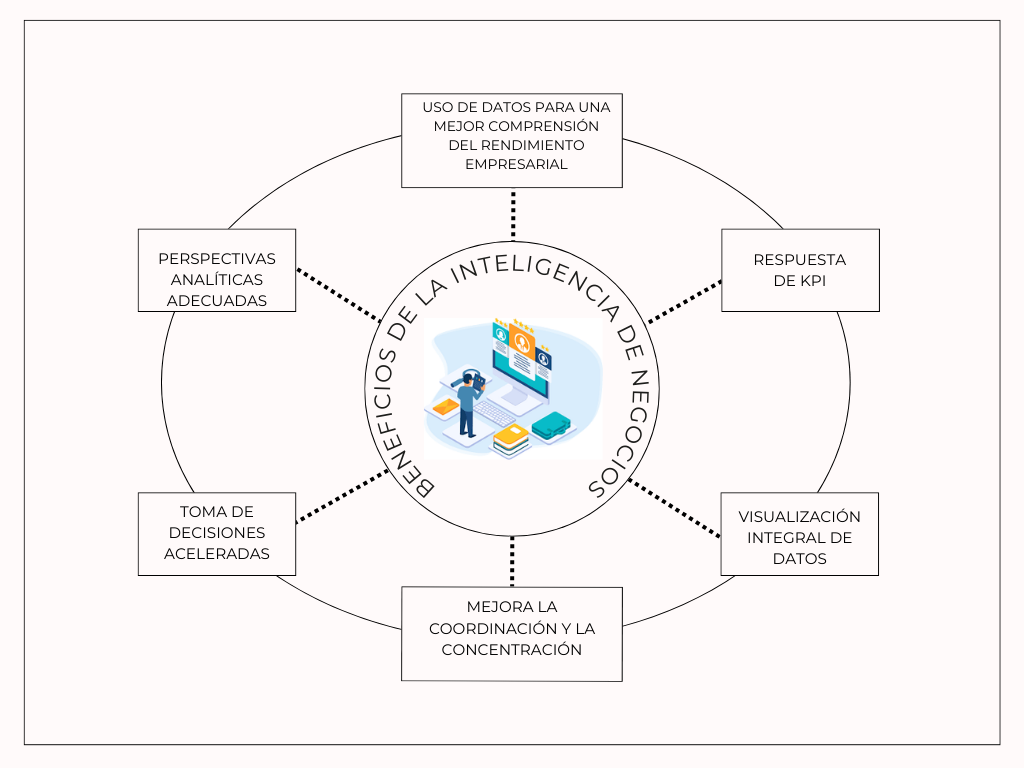
1. **Estadística y econometría:** esto lo que incluye es la teoría estadística del reconocimiento de imágenes, también de los métodos taxonométricos y de pronóstico, entre otros más (Slusarczyk Antosz, 2024).
2. **Investigación operativa:** aquí lo que hace referencia es a la programación lineal, teoría de decisiones y de juegos (Slusarczyk Antosz, 2024).
3. **Inteligencia artificial:**  en este punto lo que incluye son métodos de búsqueda heurística, aprendizaje automático, sistemas expertos, algoritmos genéticos, redes neuronales artificiales y sistemas de razonamiento analógico (Slusarczyk Antosz, 2024).
4. **Tecnologías de bases de datos:** incluye modelado de datos, lenguajes de consulta, optimización de consultas y métodos de indexación

#### Componentes de la rama inteligencia de negocios

**Tabla 7.** Componentes de la rama inteligencia de negocios

|  |  |  |
| --- | --- | --- |
| Componente | Acción | Aplicación |
| Proceso interactivo | La inteligencia de negocios la acción que facilita realizar es recopilar, limpiar y examinar información de manera continua, en lugar de hacerlo solo en un momento concretado. | Su aplicación, lo que permite es integrar, limpiar y transformar los datos y estos ponerlos en conocimiento. |
| Explorar | La inteligencia de negocios en este componente es aquel que permite entender lo que pasa en la empresa, emprendimiento, organización, y además en varios sectores ya sea de la salud, del medio ambiente entre otros. Con esto lo que nos permite es poder acceder a información que nos facilitará al momento de realizar la interpretación de los sucesos. | Identificar riesgos y oportunidades. |
| Analizar | La inteligencia de negocios busca descubrir las variables, sus relaciones y las probabilidades de que se repitan o cambien dichas relaciones. | Proveer información clave para tomar decisiones acertadas en un entorno incierto. |
| Data warehouse | La inteligencia de negocios se sustenta en el uso práctico de los datos almacenados en los sistemas. | Permite almacenar y disponer herramientas para análisis de datos y administración de contenidos. |
| Tecnología | La inteligencia de negocios tiene como propósito un objetivo específico. | Respaldar las decisiones tomadas tanto internamente como externamente. |
| Comunicar | La inteligencia de negocios lo que permite es poder transmitir lo que se ha encontrado por los datos recolectados y esto a su vez ayuda a realizar los ajustes necesarios en la organización para mejorar su competitividad y obtener una ventaja estratégica. | Diseñar un plan estratégico adaptado a las necesidades particulares de cada empresa. |

*Información tomada de (García-Jiménez et al., 2021). Elaborado por los autores*.



**Figura 6.** Beneficios de Inteligencia Artificial. Información tomada de (Sanchez Lennini, 2024). Elaborado por los autores.

**Herramientas del Business Intelligence**

**Tabla 8.** Herramientas del Businees Intelligence

|  |  |
| --- | --- |
| Herramientas | Descripción |
| Generadores de informes | Se emplean para elaborar informes estandarizados que contribuyen a grupos más extensos, como departamentos o toda la empresa |
| Consultas e informes | Son empleadas por los usuarios finales para desarrollar informes adaptados para su empleo particular. |
| Herramienta OLAP | Este instrumento permite a los usuarios manipular grandes volúmenes de datos usando estructuras multidimensionales. Esto facilita a la exploración de la información desde múltiples puntos de vista y a través de diferentes lapsos de tiempo, otorgando una apreciación más profunda. |
| Herramientas de Dashboard y scorecard | Esta herramienta distribuye una representación instantánea de la información esencial. Los usuarios tienen la capacidad de acceder a datos críticos con solo un clic y estudiar detalladamente para observar los detalles de los informes. |
| Herramientas de planificación, modelización y consolidación | Esta herramienta ayuda a implementar estrategias y llevar a cabo simulaciones basadas en la información que está presente. |
| Herramientas de data mining | Estos recursos contribuyen a los usuarios a edificar modelos estadísticos que reflejan las actividades del negocio. El data mining es imprescindible para detectar patrones ocultos en los datos, lo que a su vez brinda la posibilidad de resolver problemas y tomar decisiones decisivas. |

*Información tomada de (Casanoves, 2021). Elaborado por los autores*.



**Figura 7.** Herramientas del BI. Tomada de (Sign Consulting, 2023).

**Diferencias claves entre Inteligencia de Negocio, Analítica de Negocio y Big Data**

**Tabla 9**. Diferencias Claves

|  |  |  |  |
| --- | --- | --- | --- |
| Característica | Inteligencia de Negocios (BI) | Analítica de Negocio (BA) | Big Data |
| Madurez | Alta | Alta | En desarrollo |
| Herramientas | Consultas, sistemas de alertas, informes (reporting), OLAP, entre otros. | Clasificación, agrupamiento (clustering), regresión, etc. | Aprendizaje automático (Machine Learning), Aprendizaje profundo (Deep Learning), visualización de datos, etc. |
| Enfoque | Responde a preguntas como: ¿Qué ocurrió y cómo?, ¿cuánto?, ¿con qué frecuencia?, ¿cuál es el problema?, ¿qué acciones se requieren? | Explora el “porqué” de los eventos, ¿qué sucederá si las tendencias continúan?, ¿qué podría pasar después?, ¿cuál es la mejor opción futura? | Se centra en la recolección, el almacenamiento, el procesamiento y el análisis de grandes volúmenes de datos. |
| Uso | Principalmente reactivo (responde a eventos pasados) | Predictivo (anticipa), Proactivo (actúa con anticipación), Prescriptivo (recomienda acciones) | Abarca todas las modalidades de uso de las categorías anteriores. |
| Tipo de dato | Predominantemente estructurados | Estructurados y semi-estructurados | Cualquier tipo de dato, con un énfasis en los no estructurados. |
| Dificultad en el manejo de los datos | Baja | Baja a media | Alta |
| Alcance | Orientado a la alta dirección (nivel estratégico) | Centrado en los procesos de negocio | Extendido (vertical y horizontal) / Abarca procesos específicos. |

*Información tomada de (Curto, 2020). Elaborado por los autores*.

### Arquitectura tecnológica

La arquitectura está compuesta por varios elementos que cooperan entre ellos para recolectar, producir y a su vez poder mostrar los datos, a continuación, se nombrarán los diferentes componentes que son parte de este proceso.

**Fuentes de datos**

Esto hace referencia a los diferentes archivos que se utilizarán para este proceso. Se pueden incluir archivos .csv, archivos json, archivos txt, entre otros más que existen.

**Extracción de datos**

Este componente lo que hace es recopilar los datos de las fuentes escogidas y así guardarlas para realizar el procesamiento, esto hace referencia a la transformación.

**Transformación de datos**

Su principal función es poder identificar que datos presentan irregularidades y así poder corregir los errores que se presentaron al momento de la extracción de datos. Las irregularidades que podemos mencionar son: datos repetidos, valores nulos o datos que no se van a utilizar.

**Carga de datos**

El principal objetivo es poder almacenar los datos una vez estén transformado, para así poder analizarlos posteriormente.

**Modelo de datos**

En este componente se realizará el diseño del modelamiento de datos utilizando el modelo de estrella o copo de nieve.

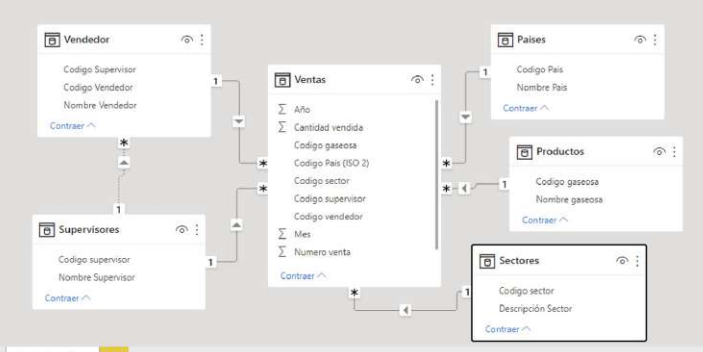
**Diseño de copo de nieve y estrella**

Según (Sarango et al., 2025) :

El diseño de estos esquemas es un elemento primordial del modelado de datos en BI, puesto que especifica su estructura y la relación entre entidades contribuyendo a una interpretación eficiente.

El esquema estrella distribuye los datos en una arquitectura central alrededor de una tabla de hechos. Este diseño es fácil de aplicar lo que permite realizar consultas inmediatas y fácil de ajustar.

El esquema copo de nieve es una alternativa más normalizada del esquema estrella. En vez de asociar de forma directa las dimensiones a la tabla de hechos, las tablas de dimensiones están segmentadas en subdimensiones extras para erradicar la redundancia.



**Figura 8.** Modelado de datos. Tomado de (Slusarczyk Antosz, 2024).

**Análisis y visualización**

Después de haber cargado los datos obtenidos en el proceso anterior y se haya realizado el modelado correctamente. Es aquí donde podremos analizar y visualizar mediante herramientas que permiten la visualización como Tableau, Looker Studio y Power BI. Esto es lo que permitirá a los usuarios identificar diversos patrones, tendencias y con esto podrán tener información real y necesaria para la toma de decisiones.

Interfaz de usuario gráfica, Aplicación, Teams

El contenido generado por IA puede ser incorrecto.

**Figura 9**.Arquitectura del sistema. Elaborado por los autores.

#### Fuentes

La información que se va a recuperar es sobre las noticias de Ecuador de los siguientes orígenes de datos:

* Google Fact Check Tools
* Corpus públicos
* La red social X:
* Ecuador Chequea
* Ecuador Verifica

Lo primero que se efectuará es la extracción de datos de las fuentes como Google Fact Check Tools, Corpus públicos y la red social X de Ecuador Chequea y Ecuador Verifica. Este proceso se lo implementará por medio del web Scraping, esto permite obtener los datos de manera directamente desde Google Fact Check Tools y de la red social X. Se utilizará varias librerías para poder realizar la extracción de los datos entre ellos tenemos, selenium, ChromeDriverManager y pytesseract y Python.

#### Extracción, Transformación y Carga

El proceso ETL en este proyecto, tiene como propósito procesar datos de las diferentes fuentes que se van a obtener las noticias para luego poder visualizarlos en un dashboard.

Este proceso tiene varias fases a continuación las mencionaremos:

**Extracción:**

Una vez que se realiza la extracción previa de cada fuente oficial, esta será almacenada en una colección con los datos crudos, posteriormente se debe extraer estos datos para proceder a la transformación.

**Transformación:**

Durante esta etapa, se llevará a cabo la obtención de datos limpios para ellos debemos realizar:

* Eliminar datos que tengan el id repetido o que el título aparezca más de una vez.
* Al momento que vamos a integrar las diferentes fuentes debemos de tener datos coherentes y que tengan una misma estructura.

**Carga:**

En esta sección, se guardarán los datos en la base de datos NoSQL y en este caso será en la de MongoDB.

#### Almacenamiento de datos e información

Los componentes esenciales de las fuentes de datos en la Inteligencia de negocios se detallan a continuación:

**Datawarehouse:** Conocido como un almacén de datos, tiene como finalidad hacer posible y proveer asistencias a las tareas de inteligencia de negocios. La gran cantidad de datawarehouses se han desarrollado para tareas de consulta y evaluación, por esta razón, acumulan amplia cantidad de datos la cual provienen de datos históricos (Reyes Mena, 2023).

Oracle describe lo siguiente: En virtud de su capacidad de análisis, las organizaciones tienen la capacidad de obtener información invaluable a partir de los datos y así perfeccionar las decisiones (2025).

Principales funciones que ofrece el datawarehouse según (Cruz & Muñiz, 2022):

* Facilita el almacenamiento de grandes volúmenes de datos en una base centralizada y bajo un formato estandarizado.
* Posibilita la integración y estandarización de diversas fuentes
* Hace posible tener datos históricos, ya que tiene la capacidad de guardar datos de meses o incluso de años anteriores.
* Asegura la seguridad de los datos conservarlos en un único lugar.

**Datamarts:** Muchas veces son creados y gestionados por un solo departamento, por lo general recolectan datos exclusivamente de un reducido número de fuentes, que serían capaces de ser sistemas operativos internos, un almacenamiento de datos centralizado o fuentes externas (Cruz & Muñiz, 2022).

**Tipos de Datamarts**

Están clasificados en tres categorías:

* **Datamart dependiente:** Este datamart permite consolidar toda la información empresarial en un único repositorio de datos, esto facilita a la centralización. Cuando se necesita crear uno o varios datamarts, lo mejor es desarrollarlos de manera dependientes para así asegurar la coherencia e integración entre los diferentes sistemas de almacenamiento(Cruz & Muñiz, 2022).
* **Datamart independiente:** El datamart se puede implementar sin tener la necesidad de hacer uso del almacén de datos central. Se sugiere emplear este datamart a departamentos o grupos pequeños que existan en una organización (Cruz & Muñiz, 2022).
* **Datamart híbrido:** Permite la integración datos que originan de distintas fuentes como de un almacén de datos. Este datamart combina tantas características de los datamarts dependientes e independientes (Cruz & Muñiz, 2022).

**Proceso de carga de datos en el Data Mart**

**Carga inicial de la base de datos**

Una vez efectuado el web scraping de las tres fuentes de información (Google Fact Check Tools, Corpus y la red social X), los datos recolectados se registran en la base de datos de MongoDB. Esta fase es fundamental, porque ofrece la capacidad de tener un repositorio intermedio en el que los datos son facultados para ser comprobados y ajustados.

**Generación de un job**

Cuando los datos estén cargados en MongoDB, se construye un job el cual gestionará de la transformación de los datos. El job se encargará de realizar el proceso automático que se implementa en un entorno ETL.

**Transformación de datos**

En esta fase, lleva a cabo la realización de tareas como la limpieza, estandarización y la optimización de los datos. Garantiza que los datos permanezcan en un formato idóneo y de excelente calidad.

**Modelado de Datos**

Se confecciona un modelo en concreto para cada fuente de datos que se represente la estructura y las relaciones de los datos.

**Carga en el Data Mart**

Tras la obtención de los datos transformados y modelados, se pone en marcha la carga en el data mart. Esta fase involucra la inserción de los datos a las tablas o colecciones respectivos.

El job que se elaboró se ejecutará para desarrollar el proceso de carga de datos. Permitirá la conexión con la base de datos NoSQL MongoDB para obtener los datos reestructurados e inserta los datos en las tablas del Data Mart según el modelo estructurado para cada fuente.

Cuando se haya completado la carga, se efectúa la revisión para confirmar que los datos se han insertado de manera apropiada y que la integridad de los datos se hay conservado.

#### Herramientas tecnológicas para extracción de datos

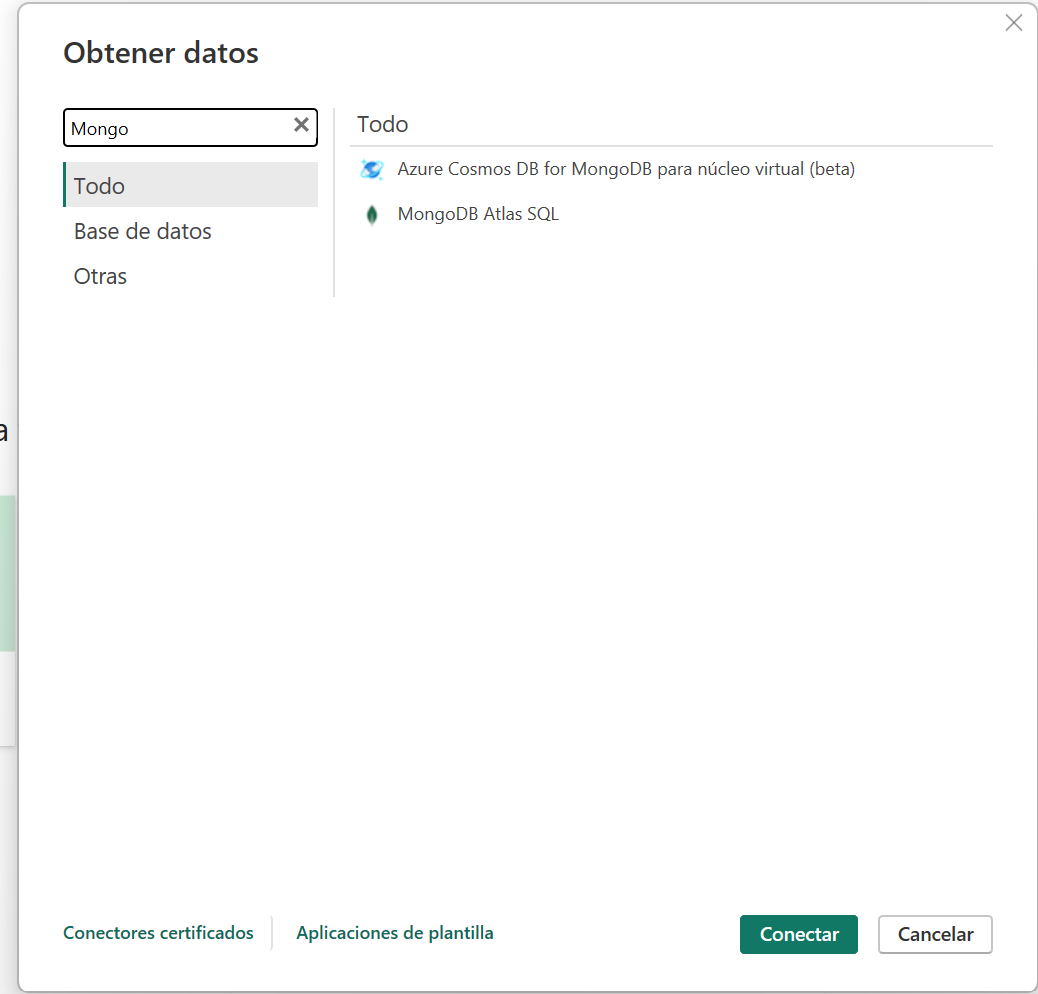
**Power Query**

Según (Microsoft Learn, 2025)

Power Query sirve como un motor para la preparación y transformación de datos. Ofrece una interfaz gráfica para la adquisición de datos desde diversas fuentes y un editor que está especializado para poder aplicar transformaciones. La flexibilidad de este motor permite su uso en múltiples productos y servicios, lo que determina que el destino de los datos varíe según la aplicación de Power Query. Con Power Query, es posible llevar a cabo el proceso completo de extracción, transformación y carga de datos (ETL).



***Figura 10.*** *Beneficios de Power Query. Información tomada de (Digari Dharmendra, 2024).Elaborada por los autores*

.

**Figura 11.** Conexión a la base de datos NoSQL(MongoDB Atlas) mediante la interfaz SQL de Atlas desde Power Query Desktop. Imagen capturada por los autores.

***Azure Data Factory***

Azure Data Factory (ADF) es un servicio de integración de datos basado en la nube concedido por Microsoft este ofrece una plataforma robusta para diseñar y ejecutar procesos ETL (Goel, 2023).

**Web scraping**

El web scraping alude a un sistema autónomo que obtiene datos y colecciona información de sitios web en internet. Algunas plataformas web obstaculizan que los agentes automatizados acudan a sus páginas mediante sistemas de detección de web scraping (Tarale, 2025).

La realización del web scraping se fundamenta en el modelo de objetos del documento (DOM), la visión artificial y el procesamiento del lenguaje natural por medio de técnicas que facilitan simular la navegación humana para consolidar el contenido de las páginas web para su evaluación (Tarale, 2025).

La biblioteca de solicitudes opera como herramienta preestablecida para las solicitudes HTTP en el lenguaje de programación Python.

Este sistema descarta la lógica compleja de las solicitudes HTTP a través de la API fácil de emplear, de forma que los usuarios puedan administrar los servicios mientras operan los datos (Tarale, 2025).

Selenium WebDriver trabaja como un framework web que habilita el procesamiento de pruebas en navegadores. La herramienta activa pruebas de aplicaciones web mediante sistematización, lo que asegura el rendimiento esperado (Tarale, 2025).

Para este proyecto utilizaremos el web scraping para evadir costos elevados de las APIs de muchas plataformas. El scraping permite adquirir la misma información de forma gratuitita con tal que se cumplan los términos de servicio del sitio.

Con el scraping, por más que la estructura HTML de la página se modifique, es más sencillo ajustarse a los selectores o mejorando la lógica de extracción.

#### Herramientas tecnológicas para transformar datos

**Talend Open Studio**

Es una herramienta ETL open source más comunes, con una interfaz fácil de usar y un amplio grupo de componentes de integración de datos. Permite el tratamiento por lotes, flujo de datos en tiempo real y pipelines de datos en la nube. La versión gratuita tiene ciertas limitaciones, debido que algunas opciones avanzadas solo es la edición empresarial (Leo & James, 2024).

**Apache NiFi**

Es reconocida por su potencial para administrar el fujo de datos en tiempo real. Su interfaz web facilita a los usuarios desarrollar y controlar de manera fluida las canalizaciones de datos.

Idóneo para la integración de datos en streaming, lo que transforma en una magnífica opción para aplicaciones de Internet de las cosas (IoT) y big data(Leo & James, 2024).

**Pentaho Data Integration**

Brinda un robusto conjunto de funcionalidades ETL, abarcando la compatibilidad con transformaciones avanzadas de datos, procesamiento paralelo y amplias opciones de conectividad. Sobresale por su versatilidad en el manejo de datos estructurados y no estructurados (Leo & James, 2024).

**Job**

La herramienta tecnológica de transformación que se utilizará en este proyecto es un job que permitirá realizar transformaciones específicas antes que se analicen los datos en Power BI. Este job incluirá la limpieza de datos, la normalización.

Además, se utiliza esta herramienta porque mantendrá actualizada las colecciones según el intervalo que se le programa para ejecutarse.

#### Herramientas tecnológicas para base de datos

**Couchbase**

Según lo indicado por Marrero

Es una base de datos distribuida que está basada en documentos. Se beneficia de una capa de almacenamiento en memoria de caché, facilita conceder operaciones de desarrollo inmediato, almacenamiento, innovación y restauración.

Couchbase permite tiempos de respuesta de rango inferior y ha sido perfeccionado para el almacenamiento de datos de gran concurrencia. El sistema es de manera progresiva debido a una arquitectura de “shared nothing” y puede optimizarse el rendimiento global agregando más nodos (2023).

**Firebase Realtime Database**

Es una base de datos que es administrada por Google y colocada en la nube. Los datos se guardan en formato json y se alinean en tiempo real con cada usuario conectado.

Dispone de una versión o plan de pago para cumplir con las demandas a gran escala. Se puede segmentar la información en diferentes situaciones de Bases de Datos dentro del mismo proyecto de Firebase (Marrero, 2023).

**MongoDB**

Es una base de Datos NoSQL documental, es un programa de código abierto, cuenta con alta disponibilidad, es competente para escalar de forma horizontal y es orientada a documentos. Creado para desarrolladores de aplicaciones actuales y registra sus datos en colecciones de documentos flexibles (Marrero, 2023).

En MongoDB existen dos formas de dividir las colecciones:

**Particionamiento basado en rangos:** particiona la colección en función de los valores del campo que se eligen. Los lotes creados son conjuntos de documentos no cubiertos y provienen por un valor máximo y otro menor de la clave seleccionada como clave de rango (Marrero, 2023).

**Particionamiento basado en hashes:** desarrolla las claves hash de los valores de un campo en específico común a todos los documentos y emplea estos hashes para producir los lotes (Marrero, 2023).



**Figura 12.** Elementos necesarios para incorporar MongoDB para vincular con herramientas de BI. Información tomada de (Ashraf Alameer, 2019).Elaborada por los autores.

#### Herramientas tecnológicas para visualizar datos

**Propósito de las herramientas de visualización**

El propósito de las herramientas de visualización es exhibir información y datos en un entorno visual. Esto simplifica la interpretación y exploración de los datos y los datos obtenidos del procesamiento. El medio visual es el sistema de percepción humano más ágil, lo que influye progresivamente que vez más personas se inclinen por observar o leerlos(Skender & Manevska, 2022).

**Características de las herramientas de visualización**

**Tabla 10**. Características de las Herramientas de Visualización

|  |  |
| --- | --- |
| Características | Descripción |
| Recopilar y gestionar datos | Facilitan la incorporación de datos procedentes de numerosas fuentes, como base de datos, hojas de cálculo, etc. |
| Análisis de datos | Proporcionan a los usuarios indagar e interpretar los datos de forma exhaustiva |
| Visualización de datos | La facultad de procesar los datos en presentaciones comprensibles |
| Dashboards personalizables | Permiten a los usuarios generar informes y dashboards personalizados que se adaptan conforme a las necesidades |
| Acceso y colaboración en tiempo real | Otorgan el acceso a los datos y la cooperación entre usuarios en tiempo actual |
| Seguridad y gobernanza de datos | Deben proporcionar que los datos sean accesibles solo por usuarios habilitados y que satisfagan las directrices de privacidad y protección de datos |

*Información tomada de(Rodríguez, 2024b). Elaborado por los autores*.

**Tipos de herramientas de visualización**

En la siguiente tabla se realiza la comparación de las herramientas de visualización de datos:

**Tabla 11.** Tipos de herramientas para la visualización de los dashboards.

|  |  |  |  |
| --- | --- | --- | --- |
| Características | Power BI | Tableau | Zoho Analytics |
| Solución | Incorpora los datos de diversas fuentes y a su vez los cambia en información consistente en un solo panel visual de gestión de negocio datos y pueden ser en diferentes formatos. | La solución que presenta es que lleva a cabo el trabajo mediante una manera más eficaz con la aplicación de mensajería para Slack | Colaboración a los usuarios empresariales a vincularse, a examinar información y generar visualizaciones de datos que resulten llamativas |
| Integración | Inclusión de Copilot en Power BI asume, en otro orden de cosas, la manifestación de la inteligencia artificial generativa ha causado que los usuarios se centren en tareas que son más importantes | Las nuevas capacidades que ofrece Data Cloud para Tableau es que agilizan que las visualizaciones se integren de forma directa en la nube de datos Data Cloud | Unifica lo que puede ser factible proporcionar datos de servicios de almacenamiento en línea como los conocidos Zoho Docs, Dropbox, Microsoft OneDrive y Google Drive |
| Server | Se utiliza para producir informes paginados que luego se suministran en el servicio de Powe BI, y un servidor de informes local en el que es viable difundir los reportes de Power BI después de haberlos creado en el Power BI Desktop. | Herramienta de análisis empresarial, por ejemplo, lo que permite es añadir almacenes de identidades locales adicionales | La realización versátil ya sea en nube pública o entornos locales, da la oportunidad de que los datos se abastezcan desde hojas de cálculo y archivos planos como por ejemplo los archivos Excel, CSV, HTML, XML, JSON |

*Información tomada de(Regina de Miguel, 2023). Elaborado por los autores*.

**Arquitectura de Power BI**

**Tabla 12.** Arquitectura de Power BI

|  |  |  |  |
| --- | --- | --- | --- |
| Componente Principal | Descripción | Características Clave |  |
| Fuentes de Datos | Power BI Desktop tiene la capacidad de conectarse a una amplia gama de fuentes de datos, tanto internas como externas, para extraer la información necesaria. | * Bases de Datos Relacionales: SQL Server, MySQL, PostgreSQL * Archivos: Excel, CSV, XML, JSON. * Servicios en la Nube: Azure, Google Analytics, Salesforce. * APIs y Servicios Web |  |
| Power Query | Es aquel que permite realizar el ETL en Power BI Desktop, facilita a los usuarios conectarse, combinar y refinar datos de diferentes fuentes antes de cargarlos en el modelo. | * Conexión con diversas fuentes de Datos: Brinda facilidad al momento de realizar conexiones con diferentes fuentes de datos * Transformación de Datos: Ofrece gran variedad de funciones para filtrar, agrupar, agregar, limpiar y transformar datos. * Combinar Datos: Permite integrar datos de diferentes fuentes en un único modelo de datos. |  |
| Modelo de Datos | Es el espacio en Power BI Desktop donde se almacenan los datos ya transformados. Es donde, los usuarios pueden definir relaciones entre las tablas y crear cálculos personalizados mediante el lenguaje DAX. | * Relaciones entre Tablas: Permite definir enlaces entre tablas para simplificar el análisis * Columnas Calculadas: Creación de elementos calculados usando DAX para análisis avanzados. |  |
| Visualización de Datos | Power BI Desktop proporciona un amplio conjunto de herramientas para que los usuarios creen informes y dashboards interactivos que representen sus datos de forma gráfica. | * Gráficos y Tablas: Incluye gráficos de barras, líneas, pastel, tablas y matrices * Mapas: Permite la visualización de datos geoespaciales * Dashboards Interactivos: Paneles que integran múltiples visualizaciones y facilitan la interacción dinámica con los datos. |  |
| Publicación y Compartición | Cuando se han terminado de desarrollar informes y dashboards en Power BI Desktop, se pueden publicar en el servicio Power BI para su distribución y colaboración con otros usuarios. | * Publicación en Power BI Service: Facilita subir informes y dashboards a la nube para compartirlos * Colaboración en tiempo Real: Ofrece herramientas para trabajar en equipo y revisar informes de manera conjunta. |  |

*Información tomada de(Rodríguez, 2024b). Elaborado por los autores*.

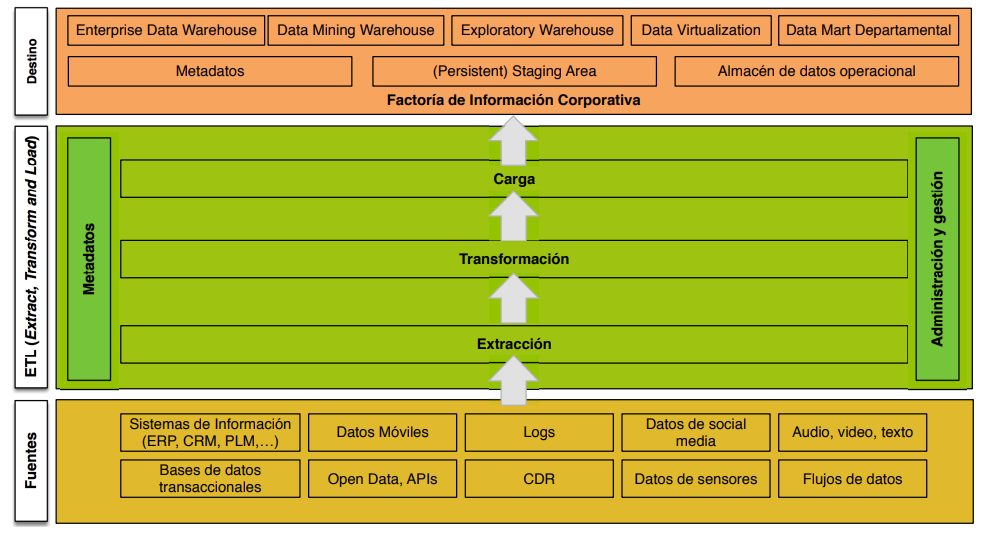


**Figura 13**. Arquitectura de Power Bi. Tomada de (Ferrer, 2020).

**Proceso ETL**

El proceso ETL (Extracción, Transformación y Carga) implica recopilar datos desde múltiples fuentes, modificar su estructura o contenido según necesidades específicas, y almacenar los resultados en un sistema de datos. Su propósito es optimizar el flujo de información, permitiendo modificar grandes cantidades de datos de forma automatizada y ágil, lo que facilita su uso posterior en análisis o procesos estratégicos (López Espinoza, 2024).

Además, lo que indica (López Espinoza, 2024) es que, aunque presenta ventajas, su principal inconveniente radica en la complejidad para gestionar y mantener el proceso en escenarios con grandes volúmenes de datos. Para mitigar esto, una estrategia efectiva es fragmentar el ETL en componentes más reducidos, lo que facilita su administración y optimización. Otro desafío relevante es el impacto negativo en el rendimiento del sistema, especialmente en las fases de carga, donde los tiempos pueden ser prolongados. Una solución viable es aplicar técnicas de procesamiento distribuido, que permiten equilibrar la carga de trabajo y acelerar las operaciones.



**Figura 14.**Proceso ETL. Información tomada de (Curto, 2020).

**Implementar DirectQuery en Power BI Desktop**

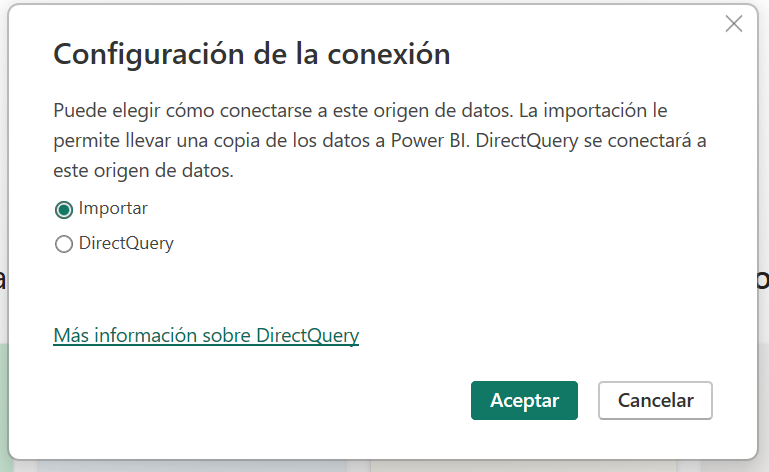
Según (Microsoft Learn, 2024a) en Power BI Desktop, no se almacenan datos locales cuando se usa DirectQuery. En su lugar, el sistema muestra tablas y columnas disponibles en el panel Datos, según el tipo de origen:

* Orígenes relacionales: Por ejemplo, lo que se menciona es SQL server. Las tablas y columnas seleccionadas se listan en el panel, pero los datos se consultan en tiempo real desde la fuente subyacente (Microsoft Learn, 2024a).
* Orígenes multidimensionales: En este punto, lo que se referencia es SAP Business Warehouse, aquí las dimensiones y medidas aparecen en el panel, reflejando la estructura del modelo de datos original (Microsoft Learn, 2024a).

Durante la creación o interacción con visualizaciones, Power BI Desktop ejecuta consultas dinámicas al origen de datos, garantizando que los resultados siempre muestren información actualizada (Microsoft Learn, 2024a).

**Implementar Importación en Power BI Desktop**

Al importar datos en Power BI Desktop, se crea una réplica local de las tablas y columnas seleccionadas. Las visualizaciones y sus interacciones se basan en esta copia almacenada. Para reflejar modificaciones en los datos originales tras la importación inicial o la última actualización, es necesario volver a importar el modelo semántico completo para sincronizar los cambios (Microsoft Learn, 2024a).



**Figura 15.** Configuración de la conexión. Imagen capturada desde Power BI por los autores.

**Visualizaciones en Power BI**

A continuación, se describen las gráficas más utilizadas para la visualización de datos

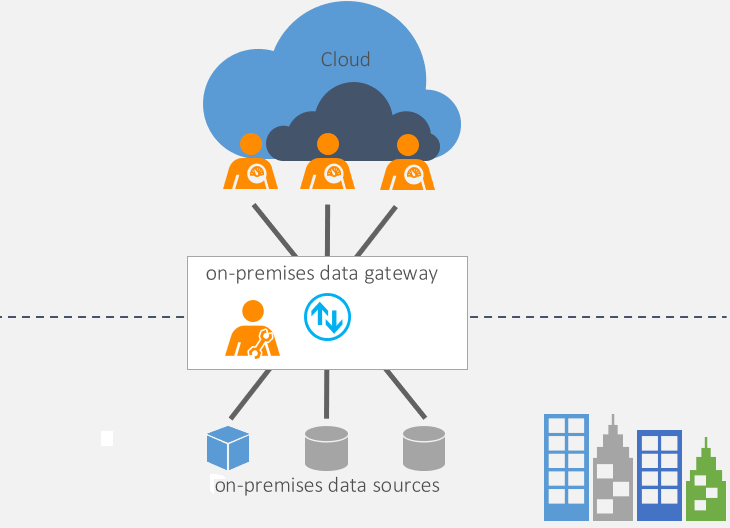
**Tabla 13.** Gráficas para la visualización de los dashboards

|  |  |  |
| --- | --- | --- |
| Tipo | Descripción | Ejemplo |
| Gráficos de áreas | Se enfatiza la magnitud del cambio con el periodo y se pueden emplear para captar la atención sobre el monto total en una orientación | Captura de pantalla de un gráfico de área. |
| Gráficos de barras y columnas | Son el patrón para indagar un valor exacto en categorías variadas | Captura de pantalla de un gráfico de barras. |
| Cards | Se presentan uno o varios datos, uno por fila | Captura de pantalla de una tarjeta de varias filas. |
| Gráficos de anillos | Presentan la relación de las partes con el todo |  |
| Número único | Se visualiza un solo hecho | Captura de pantalla de una tarjeta de número único. |
| Gráficos de líneas | Enfatizan la forma general de toda una serie de valores, usualmente a lo largo del tiempo | Captura de pantalla de un gráfico de líneas. |

*Información tomada de(Microsoft, 2024). Elaborado por los autores*.

**Power BI Gateway**

La puerta de enlace de datos local es una aplicación de software que se instala en nuestra propia red, sirviendo como un vínculo crucial. Su función principal es establecer una conexión segura y de alto rendimiento entre tus datos locales (son aquellos que residen fuera de la nube) y la suite de servicios en la nube de Microsoft. Esta puerta de enlace permite que herramientas como Power BI, Power Apps, Power Automate, Azure Analysis Services y Azure Logic Apps interactúen directamente con las bases de datos y otras fuentes de información que permanecen en tu infraestructura local. De esta manera, las organizaciones pueden seguir almacenando sus datos de forma segura en sus redes internas pueden seguir almacenando sus datos de forma segura en sus redes internas. (Microsoft, 2025).



**Figura 16**. Cómo funciona la puerta de enlace. Información tomada de (Microsoft, 2025).

**Tipos de puertas de enlace**

**Tabla 14.** Puertas de enlace

|  |  |
| --- | --- |
| Tipos | Descripción |
| Puerta de Enlace de Datos Local | Permite conectarse a múltiples usuarios a diferentes fuentes de datos locales. Con una única instalación, se puede utilizar esta puerta de enlace con todos los servicios compatibles. Es ideal para situaciones complejas donde varias personas necesitan acceder a diversas fuentes de datos. |
| Puerta de Enlace de Datos Local (modo personal) | Permite solo a un usuario a enlazarse a las fuentes de datos y no se puede compartir con otros. Una puerta de enlace de datos local (modo personal) solo se puede utilizar con Power BI. Es ideal para escenarios donde una única persona crea informes y no necesita compartir ninguna fuente de datos. |
| Puerta de Enlace de Datos de Red Virtual | Permite que varios usuarios se conecten a múltiples fuentes de datos que están protegidas mediante redes virtuales. A diferencia de las anteriores, esta puerta de enlace no requiere instalación, ya que es un servicio administrado por Microsoft. Es ideal para escenarios complejos donde varias personas acceden a múltiples fuentes de datos. |

*Información tomada de(Microsoft, 2025) por los autores*.

**Cargas de trabajo de puerta de enlace**

Según (Microsoft Learn, 2024b):

La carga de trabajo de datos almacenados en caché es un proceso para trasladar la información desde su punto de partida hasta los modelos semánticos de Power BI. Este proceso comprende de tres pasos. Como primer paso, se establece la conexión, donde la puerta de enlace de datos opera como un canal directo con la fuente de los datos, resultando importante para poner en marcha el desplazamiento de la información. En la siguiente fase, se avanza a la restauración y transformación de datos, donde la información se recupera y altera cuando sea necesario para la exploración en Power BI. En este punto, el motor de mashup de Power Query mejora el procedimiento “impulsando” las transformaciones al origen de datos, lo que es identificado como consolidación de consultas, disminuyendo la carga sobre la puerta de enlace. De todas formas, si el anidamiento de consultas no es alcanzable, la puerta de enlace desempeñará las adaptaciones, lo que conllevará un incremento en el consumo de recursos de CPU y memoria. En la etapa final, los datos procesados se trasladan al servicio Power BI, este paso demanda de una conexión a Internet estable y segura, sobre todo cuando exista grandes cantidades de datos, con el fin de garantizar que la información llegue de manera rápida y sin suspensiones, preparada para su uso en los modelos semánticos.



**Figura 17.** Puertas de enlace. Tomada de (Microsoft Learn, 2024b).

## Marco Legal

El trabajo de investigación se sustenta en los artículos que se definen a continuación:

**Ley orgánica de protección de datos personales**

**Según el reglamento de la ley orgánica de protección de datos personales:**

**Art. 1**.- **“Esta ley tiene por objeto desarrollar, proteger, promover, garantizar, regular y fomentar, el ejercicio de los derechos a la comunicación establecidos en los instrumentos de derechos humanos y en la Constitución de la República del Ecuador”**(Ley Orgánica de Comunicación, 2019)**.**

Captará la protección del derecho a desempeñar la libertad de expresión, y a explorar, adquirir y transmitir información y pensamientos de todo carácter usando los medios de comunicación.

**Art. 22.- “Derecho a recibir información de calidad. - Todas las personas tienen derecho a que la información de relevancia pública que reciben a través de los medios de comunicación sea verificada, contrastada, precisa y contextualizada.”**(Ley Orgánica de Comunicación, 2019).

**Opinión:** Este artículo da a conocer que toda información que está expuesta al público debe tener veracidad y debe de ser validada de las fuentes de donde proviene dicha información, para así evitar que a los ciudadanos le lleguen información falsa.

**Ley orgánica de transparencia y acceso a la información pública**

Según los artículos de la Ley Orgánica de Transparencia y Acceso a la Información Pública (LOTAIP):

**Art. 1.- “La presente Ley tiene por objeto garantizar y regular el derecho de acceso a la información pública en cumplimiento de la Constitución de la República del Ecuador, la ley; y, de los instrumentos internacionales ratificados por el Estado ecuatoriano.”**(Asamblea Nacional, 2023).

**Opinión:** Este artículo especifica que cualquier sujeto, sin necesidad de dar explicaciones tiene derecho a acceder a información que este en poder del Estado o de entidades públicas.

**Art. 2.- “Finalidad. La presente Ley tiene por finalidad, proteger, respetar, promover y garantizar que la información pública sea accesible, oportuna, completa y fidedigna, para el ejercicio de los derechos ciudadanos contemplados en la Constitución de la República del Ecuador y la ley.”**(Asamblea Nacional, 2023).

**Opinión:** Este artículo lo que indica es que la información pública debe ser accesible, veraz y tiene como propósito proteger, respetar y garantizar dicha información.

**Art. 7.- “Derecho de acceso a la información pública. El derecho de acceso a la información pública comprende el derecho a buscar, acceder, solicitar, investigar, difundir, recibir, copiar, analizar, reprocesar, reutilizar y redistribuir información. Toda la información producida, obtenida, adquirida, transformada o en posesión de los sujetos obligados es pública y accesible a cualquier persona en los términos y condiciones establecidos en la presente Ley, la normativa vigente y en los instrumentos internacionales aprobados y ratificados por el Estado ecuatoriano”**(Asamblea Nacional, 2023).

**Opinión:** El derecho de acceso a la información pública es un elemento clave, facilita a los ciudadanos no solo ingresar a la información que les corresponde, incluso intervenir activamente en la vida pública y en proceso decisorio.

**Normas de X**

La finalidad de X es ayudar en el debate público. Varias situaciones como la violencia, el acoso y otras maneras de actuar desmotivan a los ciudadanos a dar su punto de vista, disminuye la importancia de la conversación a nivel internacional(X Corp., 2025b).

**Autenticidad**

**Comportamientos no auténticos**

No se podrá participar en conductas de alteración de X o que impacten artificialmente el modo en que se descubre y amplifica el contenido (X Corp., 2025a).

**Manipulación de plataforma y spam**

No se puede utilizar los servicios de X de un modo que tenga el propósito de amplificar o suprimir información artificialmente, así como tampoco participar en comportamientos que manipulen o interrumpan la experiencia de las personas en X (X Corp., 2025b).

**Reglas de automatización**

* Usar tipos de automatización no basados en la API, como insertar scripts en el sitio web de X. El uso de estas técnicas puede provocar la suspensión permanente de tu cuenta (X Corp., 2017).
* No es admisible emplear ningún tipo de tarea automatizada que fomente, promueva ni incite el abuso, la violencia, el acoso o las conductas de incitación al odio, ya sea dentro o fuera de la plataforma (X Corp., 2017).

**Capítulo II**

**Metodología**

1. 1. **Tipo de investigación**

La presente investigación tiene como objetivo analizar y demostrar cómo se exterioriza la desinformación en ciertos medios, y exponer esos datos mediante una representación gráfica para proporcionar su comprensión.

Esta investigación se contextualiza de tipo descriptivo, facilitando el desglose de las características sin involucrarse en el suceso, colaborando en la construcción de visualizaciones que desplieguen de manera transparente y fácil de entender. Esta estimación es relevante para captar el comportamiento y alcance de la desinformación.

* + 1. ***Investigación descriptiva***

Según lo detallado por José Niño y María Mendoza (Niño Montero José Segundo & Mendoza Hidalgo Mary Liz, 2021):

Este tipo de investigación recibe también el nombre de investigación estadística ya que especifica los fenómenos en su forma actual. Es aplicado para clasificar y adquirir información sobre los detalles de un tema en específico.

La investigación descriptiva no está en condiciones de ponerse en práctica para generar una relación causal, en la que una variable incida en la otra.

* 1. **Enfoque de la investigación**

La presente investigación incorpora un enfoque mixto, fusionando elementos cuantitativos y cualitativos para analizar los patrones de desinformación.

* + 1. ***Enfoque cuantitativo***

El enfoque cuantitativo de esta investigación se centra en la recopilación y análisis de datos numéricos provenientes de las fuentes oficiales, estructurándolos en bases de datos para generar visualizaciones cuantitativas en Power BI, tales como gráficos de líneas, barras y mapas de calor, además del cálculo de KPIs relevantes, aplicando técnicas estadísticas descriptivas para identificar patrones y tendencias numéricas, presentando los resultados mediante gráficos, tablas y KPIs, y basando las conclusiones exclusivamente en la interpretación de los datos cuantitativos, asegurando precisión y objetividad, aunque con la limitación de excluir análisis cualitativos.

* + 1. ***Enfoque cualitativo***

Se examinará las narrativas e interpretaciones respecto a la desinformación, mediante estudios de caso y análisis de contenido. Donde se profundizará cómo las personas identifican y reaccionan a la desinformación.

* 1. **Técnicas de Recolección de datos**

La recolección de datos es valorada como un requisito previo para adquirir el conocimiento científico (Hernandez Mendoza & Duana Avila, 2020).

Todo instrumento implementado en la recolección de datos debe ser fidedigno, objetivo y legítimo, si cualquier elemento de estos no se lleva a cabo no será eficaz y los resultados logrados no tendrán validez(Hernandez Mendoza & Duana Avila, 2020).

* + 1. ***Entrevista***

**Tabla 15.** *Entrevista realizada a Ing. Elvis de la Rosa*

|  |  |
| --- | --- |
| **Ficha de Entrevista** | |
| **Fecha:** Jueves 22 de Mayo del 2025  **Lugar:** Vía zoom  **Entrevistadores:**   * Lorena Barros * Paola Barros | |
| **Entrevistado:** Elvis de la Rosa | **Cargo:** Ingeniero en Sistemas de Información |
| **Aporte Cualitativo:** En esta entrevista los temas que se abordo fue sobre las herramientas que utilizaron, que no esta funcionando en el sistema actual: web Scraping y el api de la red social X. Además, nos dio a conocer que microservicios utilizan el aplicativo. | |

*Información realizada por los autores.*

**Tabla 16.** *Resumen de la Entrevista realizada al Ing. Elvis de la Rosa*

|  |  |
| --- | --- |
| **Resumen de la Entrevista** | |
| **Fecha:** Jueves 22 de Mayo del 2025  **Lugar:** Vía zoom  **Entrevistadores:**   * Lorena Barros * Paola Barros | |
| **Entrevistado:** Elvis de la Rosa | **Cargo:** Ingeniero en Sistemas de Información |
| En esta entrevista se abordó varios temas el principal de donde recolectaban la data, donde era almacenada esta data, como era su estructura, que colecciones existen y que base de datos utilizaban, nos indicó que utilizaban base de datos NoSql (MongoDB). Además se abordo el tema de donde estaba alojado este aplicativo se nos dio a conocer que en Amazon y utilizaron el EC2 y también utilizaron Docker para poder levantar estos microservicios en la nube. | |

*Información realizada por los autores.*

* + 1. ***Revisión Documental***

Se efectuará una evaluación exhaustiva de Ecuador Chequea y Ecuador Verifica en la plataforma X ya que están dedicada a la comprobación de información. Está técnica ayudará a identificar y revisar la validez de las verificaciones de las noticias publicadas por estas plataformas.

* + 1. ***Extracción de Datos***

Se desarrollará la extracción de datos significativos sobre las verificaciones de noticias. Esto englobará la recolección de los datos en lo que concierne a la frecuencia de información malintencionada, las categorías de las noticias validadas y los orígenes de la información, lo que contribuirá a la detección de tendencias y patrones en la difusión de la desinformación.

* + 1. ***Instrumentos*** 
       1. ***Extracción de Datos implementado Web Scraping***

Se aplicará herramientas de scraping para mecanizar la obtención de los datos desde los sitios web de Ecuador Chequea y Ecuador Verifica. Este instrumento hará posible la extracción de grandes volúmenes de datos de manera eficiente.

* 1. **Estudio del proceso planteado**

El estudio del proceso planteado en la investigación describirá de forma entendible y sistematizada cómo se desarrollará la investigación desde el inicio hasta culminación. Esto incluye desde la extracción de datos, el análisis y visualización de los datos.

* + 1. ***Modelamiento del proceso***

Interfaz de usuario gráfica, Aplicación

El contenido generado por IA puede ser incorrecto.

***Figura 18****. Modelamiento de proceso propuesto. Realizado por los autores.*

* 1. **Especificaciones funcionales**

De acuerdo con lo planteado en el proyecto se especificó las siguientes funcionalidades:

* + 1. ***Historia de usuarios***

En las historias de usuarios se describirán las funcionalidades que tendrá el producto o el software según las necesidades.

Se realizará historias de usuarios para cada evento que se realizará en las dos fuentes oficiales Google Fact Check Tools, corpus y la red Social X.

**Tabla 17.** *Web Scraping de Google Fact Check Tools*

|  |  |
| --- | --- |
| **Historia de Usuario** | |
| **Código:** HU-001 | **Rol:** Administrador |
| **Alías:** Web Scraping Google Fact Check Tools | |
| **Prioridad:** Alta | **Iteración:** 1 |
| **Historia de Usuario:**  Como administradores requerimos que obtengas los datos de la página de Google Fact Check Tools | |
| **Criterios de Aceptación /Restricciones:**   * El sistema deberá recopilar los datos de Google Fact Check Tools, estos datos serán título, texto, verificador, etiquetado, fecha, link, tipo de noticia y origen * Los datos extraídos deberán ser validados para garantizar que estos datos se han recopilados de manera correcta y coherente * Los datos extraídos deberán almacenarse en la base de datos NoSQL en la colección google * Se debe tener en consideración posibles cambios en la estructura de este sitio de Google Fact Check Tools | |

*Información realizada por los autores.*

**Tabla 18***. Web Scraping de la red Social X*

|  |  |
| --- | --- |
| **Historia de Usuario** | |
| **Código:** HU-002 | **Rol:** Administrador |
| **Alías:** Web Scraping de la red Social X | |
| **Prioridad:** Alta | **Iteración:** 2 |
| **Historia de Usuario:**  Como administradores requerimos que obtengas los datos de la red Social X | |
| **Criterios de Aceptación /Restricciones:**   * El sistema deberá recopilar los datos de la red x de Ecuador Chequea y Ecuador Verifica, estos datos serán título, texto, verificador, etiquetado, fecha, link, tipo de noticia, origen, id, text, type, fullname, username, timestamp, reply\_count, retweet\_count, like\_count, visits\_count, bookmark\_count, Lang, possiblySensitive, source, tag, url, images, url * Los datos extraídos deberán ser validados para garantizar que estos datos se han recopilados de manera correcta y coherente * Los datos extraídos deberán almacenarse en la base de datos NoSQL en la colección tweet * Se debe tener en consideración posibles cambios en la estructura de la red social X | |

*Información realizada por los autores.*

**Tabla 19.** *Limpieza de Datos Extraídos de Google Fact Check Tools*

|  |  |
| --- | --- |
| **Historia de Usuario** | |
| **Código:** HU-003 | **Rol:** Administrador |
| **Alías:** Limpieza de Datos Extraídos de Google Fact Check Tools | |
| **Prioridad:** Alta | **Iteración:** 3 |
| **Historia de Usuario:**  Como administradores requerimos que el sistema realice la limpieza de datos para así tener datos de calidad | |
| **Criterios de Aceptación /Restricciones:**   * Los datos que se encuentren con el título repetido se deberán de eliminar * Los datos ya limpios se deberán almacenar en la base de datos | |

*Información realizada por los autores.*

**Tabla 20.** *Limpieza de Datos Extraídos de la red Social X*

|  |  |
| --- | --- |
| **Historia de Usuario** | |
| **Código:** HU-004 | **Rol:** Administrador |
| **Alías:** Limpieza de Datos Extraídos de la red Social X | |
| **Prioridad:** Alta | **Iteración:** 4 |
| **Historia de Usuario:**  Como administradores requerimos que el sistema realice la limpieza de datos para así tener datos de calidad | |
| **Criterios de Aceptación /Restricciones:**   * Los datos que se encuentren con el título, id repetido, y cuando encuentre el etiquetado de “No\_verificado” se deberán de eliminar * Los datos ya limpios se deberán almacenar en la base de datos | |

*Información realizada por los autores.*

**Tabla 21.** *Limpieza de datos del Corpus*

|  |  |
| --- | --- |
| **Historia de Usuario** | |
| **Código:** HU-005 | **Rol:** Administrador |
| **Alías:** Limpieza de Datos del Corpus | |
| **Prioridad:** Alta | **Iteración:** 5 |
| **Historia de Usuario:**  Como administradores requerimos que el sistema realice la limpieza de datos para así tener datos de calidad | |
| **Criterios de Aceptación /Restricciones:**   * Los datos que se encuentren con el título, id repetido se deberán de eliminar * Los datos ya limpios se deberán almacenar en la base de datos | |

*Información realizada por los autores.*

**Tabla 22.** *Transformación de datos de Google Fact Check Tools*

|  |  |
| --- | --- |
| **Historia de Usuario** | |
| **Código:** HU-006 | **Rol:** Administrador |
| **Alías:** Transformación de datos de Google Fact Check Tools | |
| **Prioridad:** Alta | **Iteración:** 6 |
| **Historia de Usuario:**  Como administradores requerimos una vez que los datos se encuentren limpios, los transforme para poder cargarlos al datamart | |
| **Criterios de Aceptación /Restricciones:**   * Los datos transformados deberán cumplir con el esquema de estrella para cargarlos en el datamart Tesis | |

*Información realizada por los autores.*

**Tabla 23.** *Transformación de datos de la red Social X*

|  |  |
| --- | --- |
| **Historia de Usuario** | |
| **Código:** HU-007 | **Rol:** Administrador |
| **Alías:** Transformación de datos de la red Social X | |
| **Prioridad:** Alta | **Iteración:** 7 |
| **Historia de Usuario:**  Como administradores requerimos una vez que los datos se encuentren limpios, los transforme para poder cargarlos al datamart | |
| **Criterios de Aceptación /Restricciones:**   * Los datos transformados deberán cumplir con el esquema de estrella para cargarlos en el datamart Tesis-tweet. | |

*Información realizada por los autores.*

**Tabla 24.***Transformación de datos del Corpus*

|  |  |
| --- | --- |
| **Historia de Usuario** | |
| **Código:** HU-008 | **Rol:** Administrador |
| **Alías:** Transformación de datos del Corpus | |
| **Prioridad:** Alta | **Iteración:** 8 |
| **Historia de Usuario:**  Como administradores requerimos una vez que los datos se encuentren limpios, los transforme para poder cargarlos al datamart | |
| **Criterios de Aceptación /Restricciones:**   * Los datos transformados deberán cumplir con el esquema de estrella para cargarlos en el datamart Tesis-corpus. | |

**Tabla 25***. Integración de Power BI*

|  |  |
| --- | --- |
| **Historia de Usuario** | |
| **Código:** HU-009 | **Rol:** Administrador |
| **Alías:** Integración de Power BI | |
| **Prioridad:** Alta | **Iteración:** 9 |
| **Historia de Usuario:**  Como administradores requerimos una vez que se conecte Mongo Altlas y Power BI | |
| **Criterios de Aceptación /Restricciones:**   * Conectarnos directamente la base de datos MongoDB Atlas desde Power BI * Visualizar tablas de hechos y dimensiones | |

*Información realizada por los autores.*

**Tabla 26.** *Creación de tableros*

|  |  |
| --- | --- |
| **Historia de Usuario** | |
| **Código:** HU-010 | **Rol:** Administrador |
| **Alías:** Creación de tableros | |
| **Prioridad:** Alta | **Iteración:** 10 |
| **Historia de Usuario:**  Como usuario final, el Sistema debe filtrar por las páginas Ecuador Chequea y Ecuador Verifica y por el Corpus | |
| **Criterios de Aceptación /Restricciones:**   * Según los datos extraídos realizar la filtración por las noticias de Ecuador Chequea, Ecuador Verifica y por el Corpus. | |

*Información realizada por los autores.*

***Tabla 27.*** *Visualización de dashboard claro y conciso*

|  |  |
| --- | --- |
| **Historia de Usuario** | |
| **Código:** HU-011 | **Rol:** Usuario final |
| **Alías:** Visualización de dashboard | |
| **Prioridad:** Alta | **Iteración:** 11 |
| **Historia de Usuario:**  Como usuarios finales se desea visualizar datos relevantes de la desinformación en el Ecuador generando gráficos y tablas | |
| **Criterios de Aceptación /Restricciones:**   * Los gráficos deben ser claros y concisos. | |

*Información realizada por los autores.*

* + 1. ***Ficha de tareas***

Una vez realizada las historias de usuarios, se procede con las tarjetas de las tareas a realizar. Estas tareas están asociadas con las historias de usuarios, en esta parte se detalla cada tarea a realizar para completar la implementación del sistema.

**Tabla 28.***Tarjeta de Tarea 001*

|  |  |  |
| --- | --- | --- |
| **Tarjeta de Tarea 001** | | |
| **Historia de Usuario:**001 Web Scraping de Google Fact Check Tools | | |
| **#** | **Nombre** | **Descripción** |
| 1 | Configuración del ambiente de desarrollo | Se debe preparar el ambiente de desarrollo para codificar el web Scraping del sitio de Google Fact Check Tools usando librerías como selenium, request y otras librerías. |
| 2 | Ejecución del script del web Scraping | Ejecutar el script del Web Scraping de Google Fact Check Tools para recopilar los datos con su respectiva estructura. |
| 3 | Validar la información recopilada | Verificar la validez de estos datos extraídos mediante pruebas, comparando la información en la red Social X y Google Fact Check Tools |

*Información realizada por los autores.*

|  |  |  |
| --- | --- | --- |
| **Tarjeta de Tarea 002** | | |
| **Historia de Usuario:**002 Web Scraping de la red Social X | | |
| **#** | **Nombre** | **Descripción** |
| 1 | Configuración del ambiente de desarrollo | Se debe preparar el ambiente de desarrollo para codificar el web Scraping de la red Social X, usando librerías como selenium, pytesseract, request y otras librerías. |
| 2 | Ejecución del script del web Scraping | Ejecutar el script del Web Scraping de la red Social X y así mismo extraer los datos con su respectiva estructura. |
| 3 | Validar la información recopilada | Verificar la validez de estos datos extraídos mediante pruebas, comparando la información en la red Social X y Google Fact Check Tools |

*Información realizada por los autores.*

**Tabla 29.** *Tarjeta de Tarea 003*

|  |  |  |
| --- | --- | --- |
| **Tarjeta de Tarea 003** | | |
| **Historia de Usuario:**003 Limpieza de datos Extraídos de Google Fact Check Tools | | |
| **#** | **Nombre** | **Descripción** |
| 1 | Identificación de datos repetidos e incorrectos | Desarrollo de un script en Mongo Atlas que detecte datos repetidos e incorrectos |
| 2 | Eliminación de datos repetidos | Eliminar datos que contengan el título repetido o el id. |

**Tabla 30.** Tarjeta de Tarea 004

|  |  |  |
| --- | --- | --- |
| **Tarjeta de Tarea 004** | | |
| **Historia de Usuario:**004 Limpieza de datos Extraídos de la red Social X | | |
| **#** | **Nombre** | **Descripción** |
| 1 | Identificación de datos repetidos e incorrectos | Desarrollo de un script en Mongo Atlas que detecte datos repetidos e incorrectos |
| 2 | Eliminación de datos repetidos | Eliminar datos que contengan el título repetido, cuando encuentre en etiquetado “no\_verificado” o el id. |

*Información realizada por los autores.*

**Tabla 31.** Tarjeta Tarea 005

|  |  |  |
| --- | --- | --- |
| **Tarjeta de Tarea 005** | | |
| **Historia de Usuario:**005 Limpieza de datos del Corpus | | |
| **#** | **Nombre** | **Descripción** |
| 1 | Identificación de datos repetidos e incorrectos | Desarrollo de un script en Mongo Atlas que detecte datos repetidos e incorrectos |
| 2 | Eliminación de datos repetidos | Eliminar datos que contengan el título repetido o el id. |

*Información realizada por los autores.*

**Tabla 32.** *Tarjeta de Tarea 006*

|  |  |  |
| --- | --- | --- |
| **Tarjeta de Tarea 006** | | |
| **Historia de Usuario:**006 Transformación de Datos de la fuente de Google Fact Check Tools en Mongo Atlas | | |
| **#** | **Nombre** | **Descripción** |
| 1 | Transformación de los Datos | Desarrollo de un script en Mongo Atlas que realice la transformación de los datos de Google Fact Check Tools una vez que se haya realizado la limpieza de datos. |
| 2 | Pruebas | Se ejecuta el script en el trigger programado en Mongo Atlas para validar que transforme correctamente estos datos y a su vez se cargue de manera correcta en cada colección. |

*Información realizada por los autores.*

**Tabla 33.** *Tarjeta de Tarea 007*

|  |  |  |
| --- | --- | --- |
| **Tarjeta de Tarea 007** | | |
| **Historia de Usuario:**007 Transformación de datos de la red Social X | | |
| **#** | **Nombre** | **Descripción** |
| 1 | Transformación de los Datos | Desarrollo de un script en Mongo Atlas que realice la transformación de los datos de la fuente de la red Social X, una vez que se haya realizado la limpieza de datos. |
| 2 | Pruebas | Se ejecuta el script en el trigger programado en Mongo Atlas para validar que transforme correctamente estos datos y a su vez se cargue de manera correcta en cada colección. |

*Información realizada por los autores.*

**Tabla 34.** *Tarjeta de Tarea 008*

|  |  |  |
| --- | --- | --- |
| **Tarjeta de Tarea 008** | | |
| **Historia de Usuario:**008 Transformación de datos del corpus | | |
| **#** | **Nombre** | **Descripción** |
| 1 | Transformación de los Datos | Desarrollo de un script en Mongo Atlas que realice la transformación de los datos de los datos recopilados del corpus, una vez que se haya realizado la limpieza de datos. |
| 2 | Pruebas | Se ejecuta el script en el trigger programado en Mongo Atlas para validar que transforme correctamente estos datos y a su vez se cargue de manera correcta en cada colección. |

*Información realizada por los autores.*

**Tabla 35.***Tarjeta de Tarea 009*

|  |  |  |
| --- | --- | --- |
| **Tarjeta de Tarea 009** | | |
| **Historia de Usuario:**009 Carga de datos en la base NoSQL Mongo Atlas | | |
| **#** | **Nombre** | **Descripción** |
| 1 | Configuración de Conexión con la base de datos | Se debe crear los usuarios que van a tener acceso a la base de datos esto se lo realiza en la misma base de datos. Una vez creado se debe configurar y establecer la conexión con el sistema y la base de datos |
| 2 | Implementar autenticación de la base de datos | Se debe implementar la autenticación y acceso para Mongo Atlas |
| 3 | Evaluaciones de la carga de datos y del acceso. | Se realiza pruebas para ver la disponibilidad de los datos y así asegurar que se puedan realizar consultas. |

*Información realizada por los autores.*

**Tabla 36.** *Tarjeta de Tarea 010*

|  |  |  |
| --- | --- | --- |
| **Tarjeta de Tarea 010** | | |
| **Historia de Usuario:**009 Integración de PowerBI y Mongo Atlas | | |
| **#** | **Nombre** | **Descripción** |
| 1 | Conexión por medio de odbc | Instalar Power BI y el conector de Mongo Atlas. |
| 2 | Configuración de origen de datos | Agregar origen de datos (Mongo Atlas y se debe colocar las credenciales). |

*Información realizada por los autores.*

**Tabla 37.** *Tarjeta de Tarea 011*

|  |  |  |
| --- | --- | --- |
| **Tarjeta de Tarea 011** | | |
| **Historia de Usuario:**010 Creación de dashboard | | |
| **#** | **Nombre** | **Descripción** |
| 1 | Diseño de dashboard | Se debe diseñar los dashboard con métricas |
| 2 | Creación de características de visualización | Se debe crear funcionalidades que permita dar a conocer las métricas de las noticias extraídas de las fuentes de la red Social X, Google Fact Check Tools y Corpus. |
| 3 | Incorporación de elementos interactivos y opciones personalizables. | Desarrollar características de interactividad y personalización en los paneles de control, permitiendo modificaciones de acuerdo con las preferencias del usuario. |

*Información realizada por los autores.*

**Tabla 38.** *Tarjeta de Tarea 012*

|  |  |  |
| --- | --- | --- |
| **Tarjeta de Tarea 012** | | |
| **Historia de Usuario:**011 Visualización de dashboards | | |
| **#** | **Nombre** | **Descripción** |
| 1 | Diseñar gráficos claros y concisos | Diseñar distintos gráficos y también incluyendo segmentación, tablas y tarjetas para que los usuarios finales puedan visualizar datos de manera coherente y clara |
| 2 | Creación de características de visualización | Se debe crear funcionalidades que permita dar a conocer datos relevantes de las noticias extraídas de las fuentes de la red Social X, Google Fact Check Tools y Corpus. |
| 3 | Implementar filtrado | Implementar la funcionalidad de filtrado para que los usuarios puedan acceder a la información según la temática que se adapte a sus necesidades. |

*Información realizada por los autores.*

* + 1. ***Requerimientos funcionales***

**Tabla 39.** *Requerimientos funcionales*

|  |  |
| --- | --- |
| **Requerimientos funcionales** | |
| **Código** | **Requerimientos funcionales** |
| RF-001 | El sistema deberá ser capaz de realizar web Scraping a la página de Google Fact Check Tools y la red social x |
| RF-002 | Deberá ser capaz de realizar el etiquetado y la clasificación por tipo de Noticia |
| RF-003 | El sistema deberá ser capaz de limpiar los datos repetidos y así proporcionar datos concisos. |
| RF-004 | Cuando el sistema haya realizado la limpieza de datos, transforme estos datos para cargarlo en el datamart de cada fuente |
| RF-005 | El sistema permitirá filtrar por etiquetado, fecha, origen, verificador |
| RF-006 | El sistema deberá tener un dashboard que muestren métricas importantes |
| RF-007 | El sistema deberá integrar Mongo Atlas y Power BI |
| RF-008 | El sistema deberá hacer la limpieza de datos automáticamente cada cierto tiempo programado en el trigger |
| RF-009 | El sistema deberá hacer la transformación automáticamente cada cierto tiempo programado en el trigger. |

*Información realizada por los autores.*

* + 1. ***Requerimientos no funcionales***

**Tabla 40.** *Requerimientos no funcionales*

|  |  |
| --- | --- |
| **Requerimientos no funcionales** | |
| **Código** | **Requerimientos no funcionales** |
| RNF-001 | El sistema deberá ser capaz de manejar volumen grande de datos. |
| RNF-002 | El sistema deberá ser capaz de tener disponibilidad siempre de sus datos para tener acceso en tiempo real |
| RNF-003 | El sistema deberá presentar dashboard que sean claros y fáciles de utilizar. |

*Información realizada por los autores.*

* 1. **Metodología**

En el proceso de investigación y desarrollo de este trabajo, se implementará la metodología CRISP-DM (Cross-Industry Standard Process for Data Mining) por la naturaleza de su enfoque y su capacidad de adaptación a diversos proyectos de exploración de datos. Esta metodología se comprende de seis fases claves que impulsarán el flujo de extracción, transformación y análisis de datos.

Comprensión del Negocio: Se establecerá los objetivos del desarrollo del proyecto y los requerimientos, afianzando que la visión de la investigación esté en armonía con las necesidades definidas.

Comprensión de los Datos: Se lleva a cabo la evaluación de la información que se captarán mediante web scraping, localizando las fuentes de datos relevantes y examinando su calidad.

Preparación de los datos: Esta fase engloba la limpieza y transformación de los datos. Además, se implementará triggers programados en la base de datos Mongo Atlas para optimizar la ejecución de la limpieza, verificando que los datos sean confiables y estén listos para el análisis e interpretación.

Modelado: En esta etapa, se empleará técnicas de modelado adecuadas para obtener información invaluable de los datos tratados.

Evaluación: Se realizará la inspección los resultados recolectados para certificar que satisfacen los objetivos definidos en la fase de comprensión del negocio.

Despliegue: Finalmente, se demostrará los resultados derivados en el análisis realizado, corroborando que sean asequibles y eficaces para los interesados.

**Capítulo III**

# Propuesta

## Introducción

### Tema

Visualización de patrones de desinformación en Ecuador mediante un dashboard en Power BI

### Objetivos

Este proyecto implica la creación de un dashboard interactivo en Power BI que permita el análisis y visualización de patrones de desinformación en Ecuador, proporcionando información clave sobre la evolución y distribución de noticias verificadas en el país.

## Ambiente de desarrollo

Con el fin de desarrollar este proyecto se utilizó los siguientes entornos:

**Tabla 41.** Ambientes de desarrollo

|  |  |  |
| --- | --- | --- |
| Ambientes de Desarrollo | | |
| Entorno | Descripción | Propósito en el proyecto |
| Python | Elabora scripts | Realizar el scraping de las diferentes fuentes de información |
| Mongo Atlas | Almacenamiento NoSQL | Guardar información extraídas de los orígenes de datos |
| Power Bi desktop | Visualizar e inspeccionar datos | Elaborar dashboard donde se visualice los patrones de desinformación |
| Power Bi Service | Plataforma en la nube que ayuda a distribuir, publicar informes | Publicar y propagar los dashboards a usuarios finales |

Información realizada por los autores.

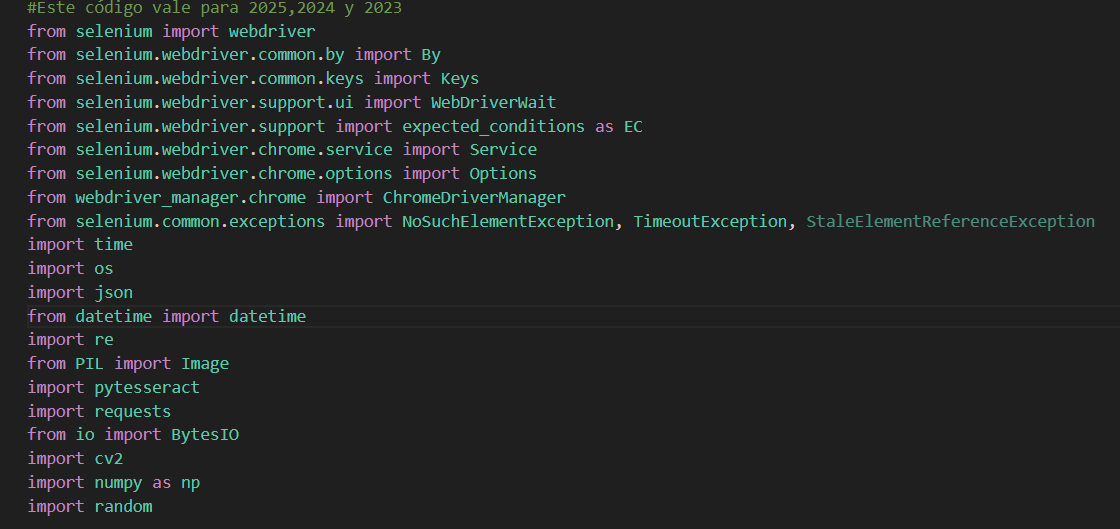
En este trabajo de visualización de patrones de desinformación de Ecuador, se elaboró scripts en la herramienta Python en la cual se realiza el web scraping automatizado para extraer la información en formato json de las fuentes de información verificadas de la red social X (EcuadorChequea y EcuadorVerifica) y de Gooogle Fact Check Tools.

En MongoAltlas se cargan los datos y se ejecutan disparadores que llevan a cabo de forma programada las tareas de limpieza, transformación, incluyendo la normalización de los campos, depuración de datos repetidos y categorización por tipo de noticia. Posteriormente, los datos transformados se cargan en las colecciones que serán utilizados para la visualización de los datos.

### Web Scraping de la red Social X

Para realizar la extracción de datos de la red social X (@EcuadorChequea), se ejecuta el script ecuador\_chequea2024.py, este script contiene librerías como selenium, webdriver\_manager.chrome, time, json, pytesseract, datetime, requests, numpy y random.

Para la cuenta de @EcuadorVerifica se dispone de otro script, que es similar al de Ecuador Chequea, lo que varía, es el código del procesamiento de imágenes, ya que estas dos cuentas tienen diferentes patrones al momento de publicar sus imágenes con los datos verificados.



**Figura 19.** Librerías utilizadas en el Web Scraping

Se debe configurar la ruta donde esta ubicado el programa que se ha instalado para el procesamiento de imágenes. Además, se muestra la lógica que se ha empleado para realizar la categorización de noticias.



**Figura 20**.Configuración de Tesseract OCR

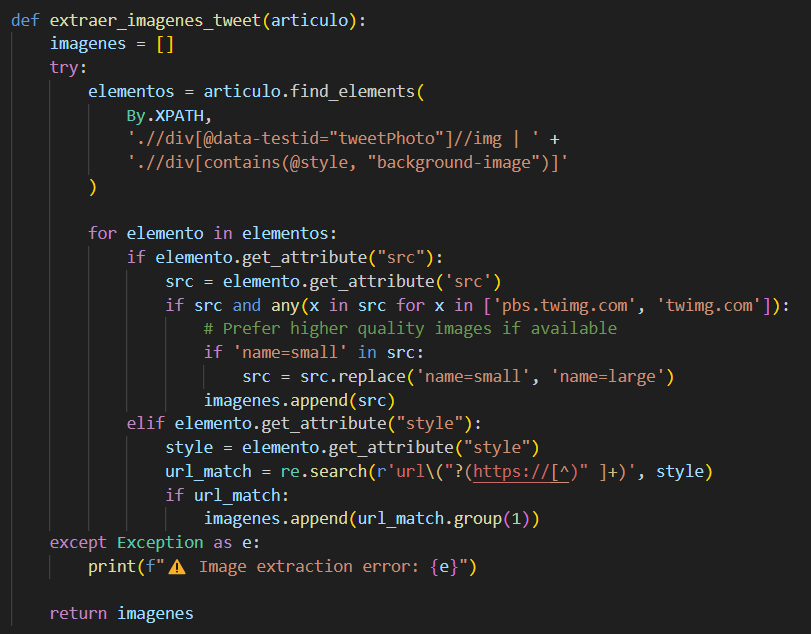
En la siguiente imagen se tienen las palabras claves que deberá detectar en las imágenes publicadas de la red Social X (@EcuadorChequea) para realizar el etiquetado de la metodología que identificamos de Ecuador Chequea que está planteado en el primer capítulo (falso, cierto, engañoso, alterado, impreciso, inverificable, sátira).

Texto

El contenido generado por IA puede ser incorrecto.

**Figura 21.** Funciones realizadas en el Web Scraping

En esta función creada extrae la imagen de las publicaciones realizadas de @EcuadorChequea



**Figura 24.** Función extraer imágenes tweet

La función **buscar\_tweets** se le debe enviar parámetros como el usuario, la fecha inicio hasta la fecha fin que desea buscar.

Texto

El contenido generado por IA puede ser incorrecto.

**Figura 31.** Buscar tweets

En esta parte se muestra la estructura que tendrá el json al momento de descargar la data



**Figura 32.** Estructura del json

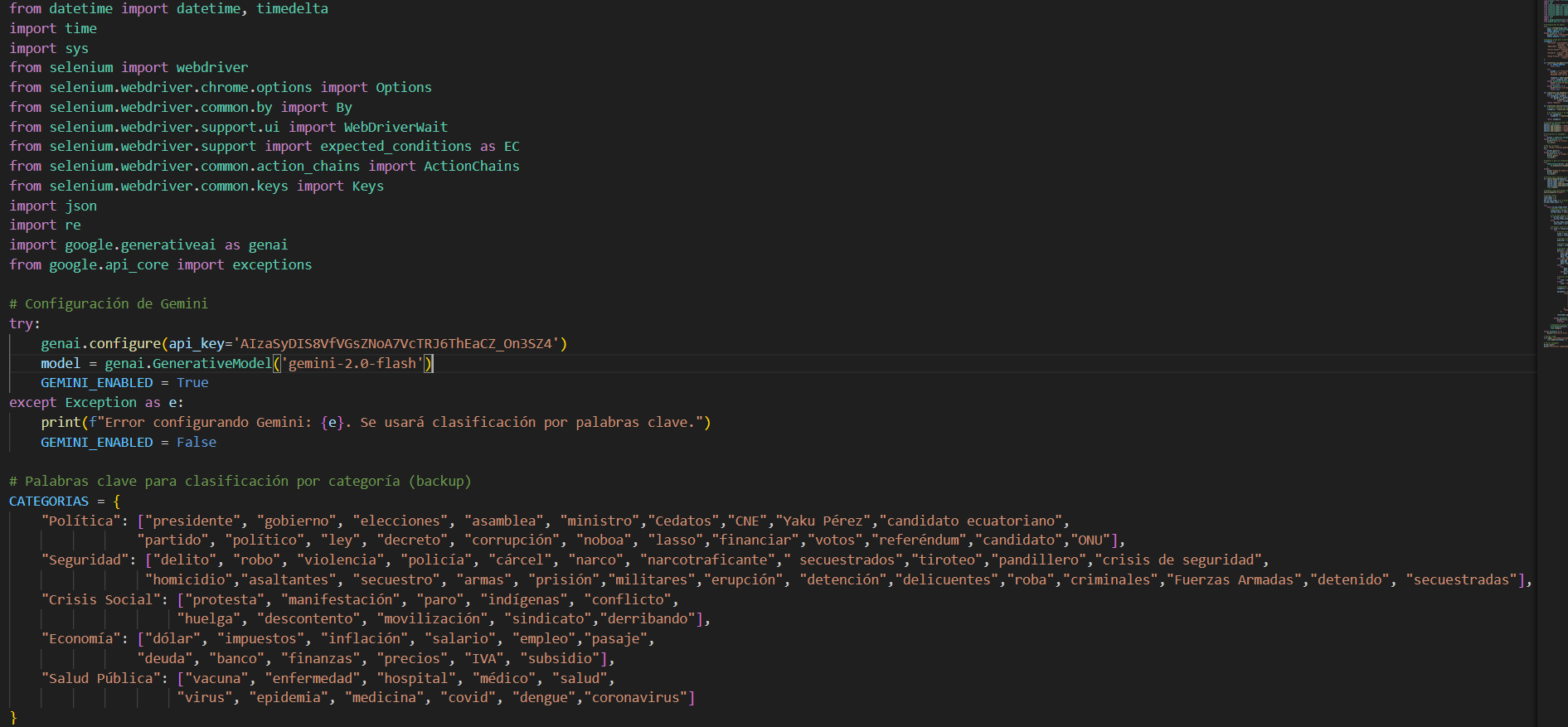
Al terminar de realizar el scraping de las fuentes de información se genera un archivo json.



**Figura 34.** Documento obtenido

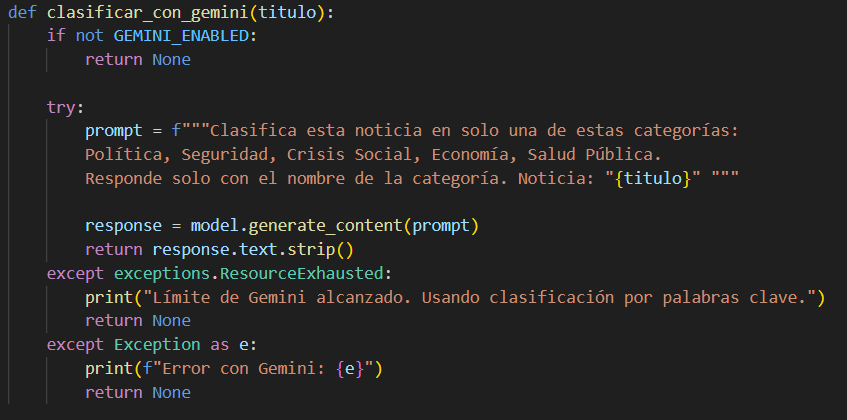
### Web Scraping de Google Fact Check Tools

Para Google Fact Check Tools se tiene otro script y se utiliza librerías que se muestran en la figura a continuación:



**Figura 35.** Librerías para el script de Google Fact Check Tools

Se tiene la función clasificar con Gemini:



**Figura 37.** Clasificar con Gemini

En el caso de que no funcione el api, se tiene una función clasificar con palabras clave, como se muestra en la imagen:

Interfaz de usuario gráfica, Texto

El contenido generado por IA puede ser incorrecto.

**Figura 38.** Clasificar con palabras clave

A continuación, se tiene la función clasificar noticias, esta función utiliza las dos funciones creadas anteriormente:

Texto

El contenido generado por IA puede ser incorrecto.

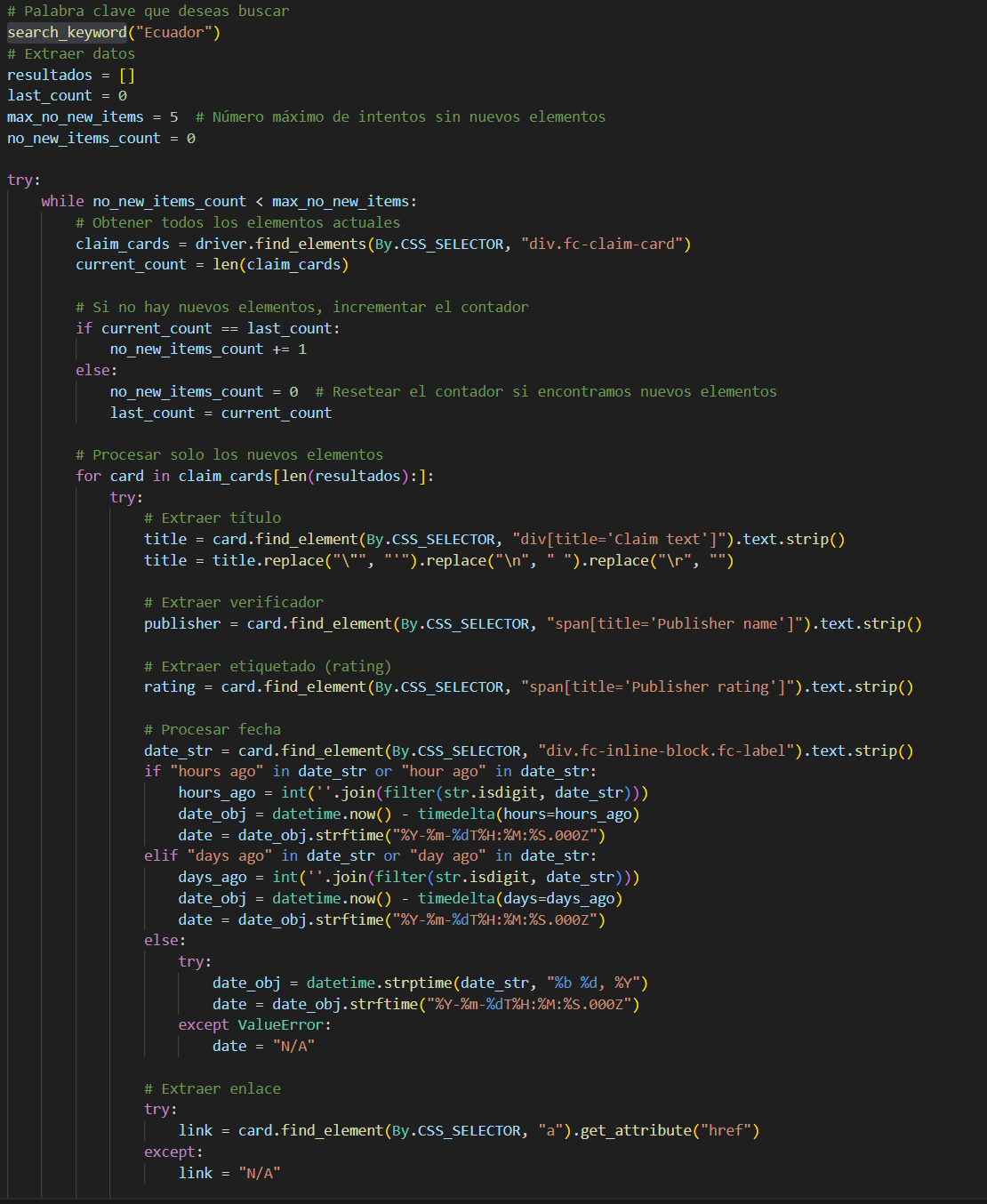
**Figura 39.** Clasificar noticia

Se configura el navegador Chrome, como se muestra en la figura:



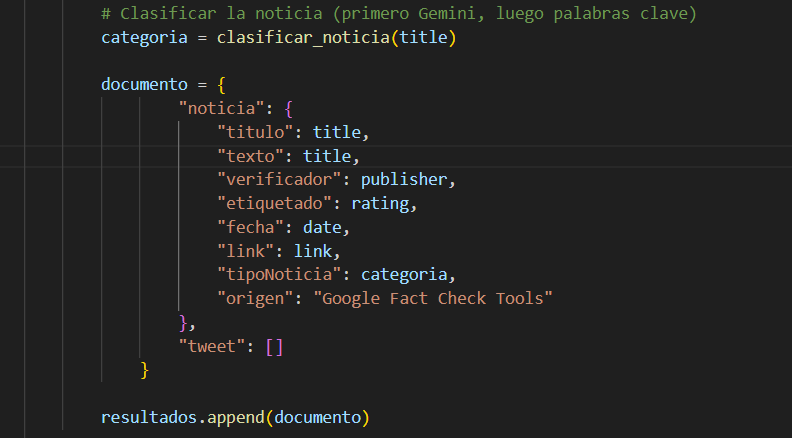
**Figura 40.** Configuración de Chrome

En esta parte del código se muestra los selectores que tiene esta página Google Fact Check Tools



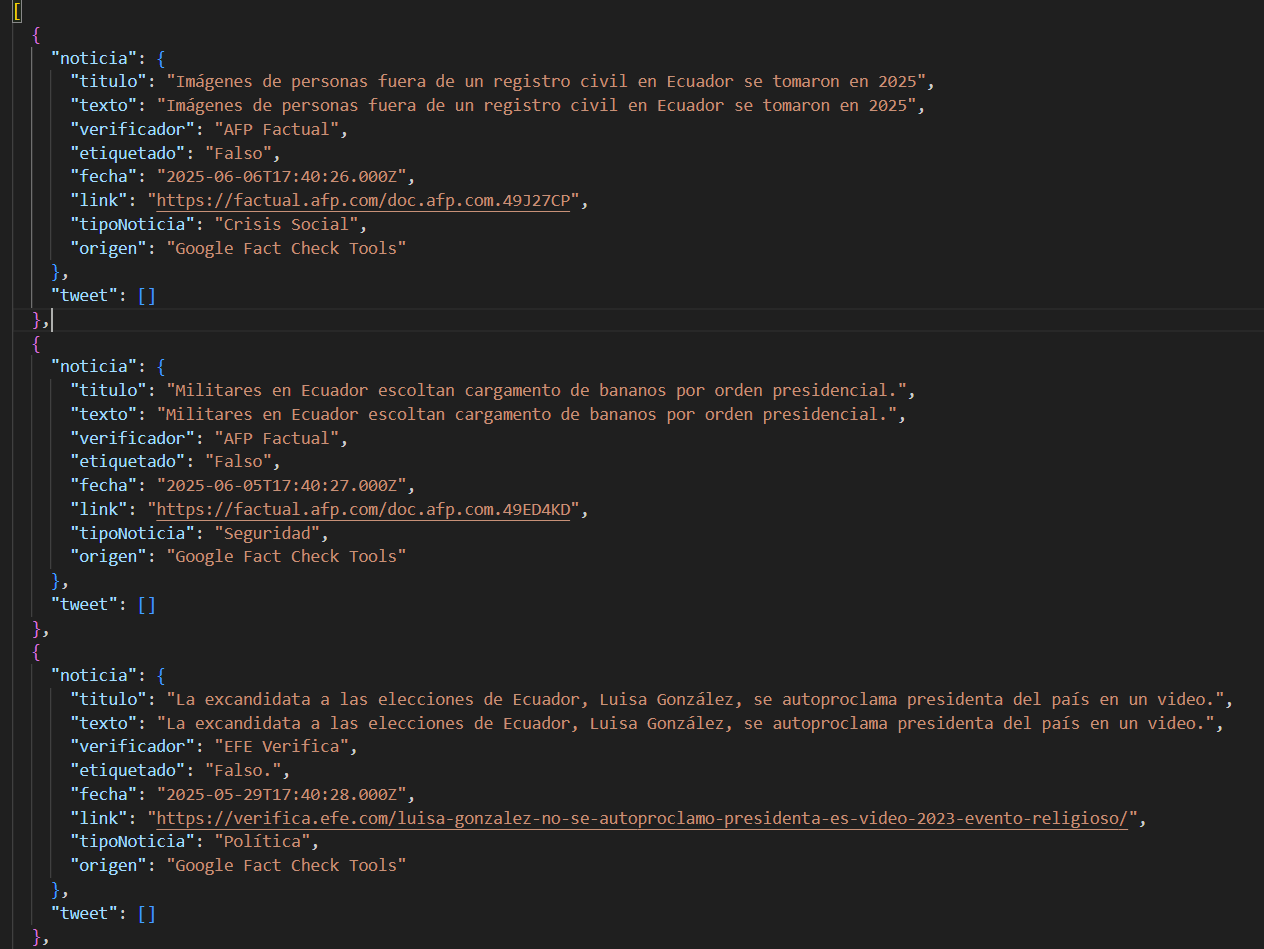
**Figura 41.** Selectores

Se muestra cómo se guardan los datos en el json



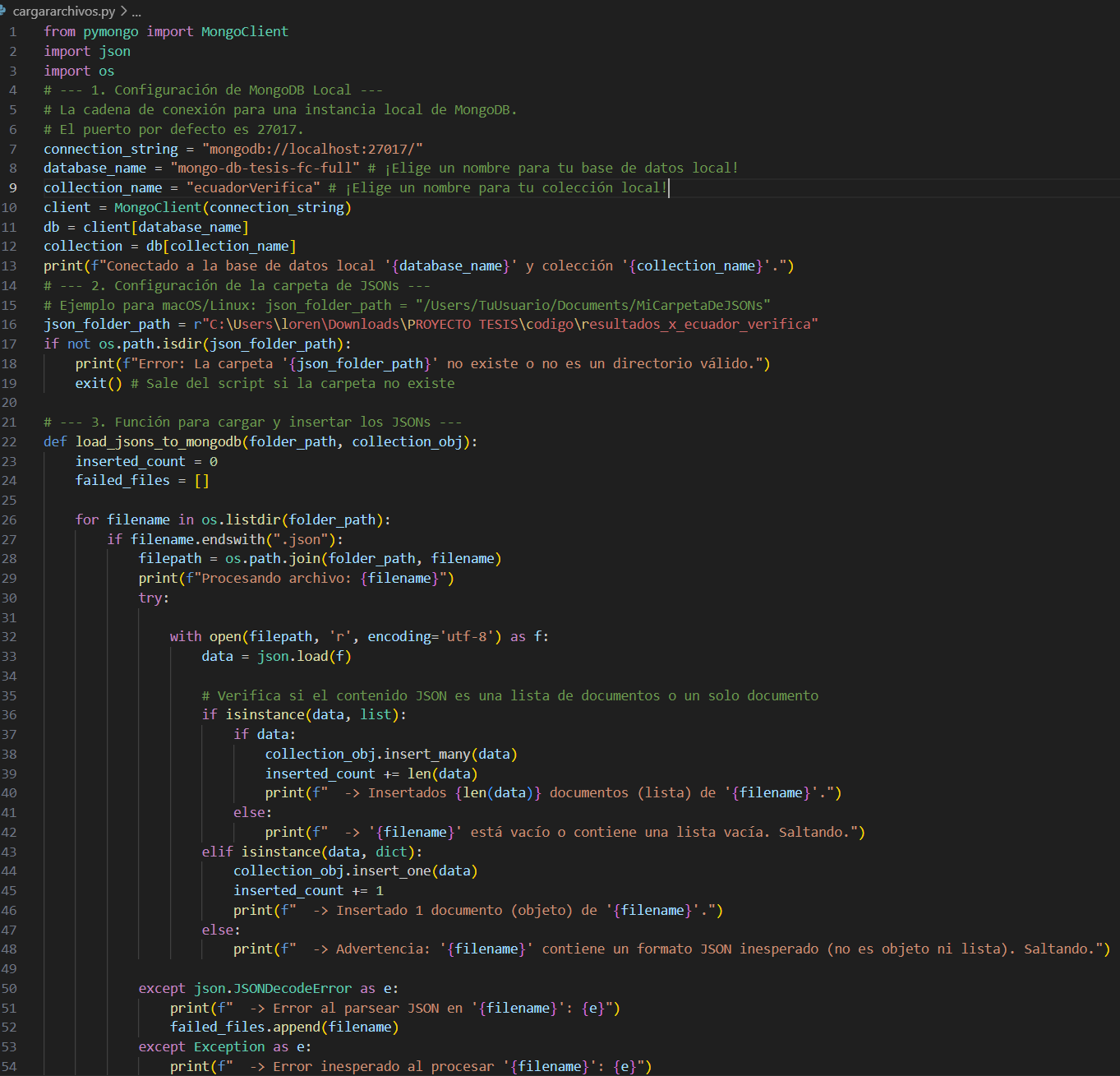
**Figura 42.** Estructura de json de los datos obtenidos de Google Fact Check Tools

Documento json generado por el script de Google\_fact.py



**Figura 43.** Documento extraído del script de Google Fact Check Tools

### Carga de datos en la base de datos local MongoDB y Migración a la base alojada en la nube Mongo Altas

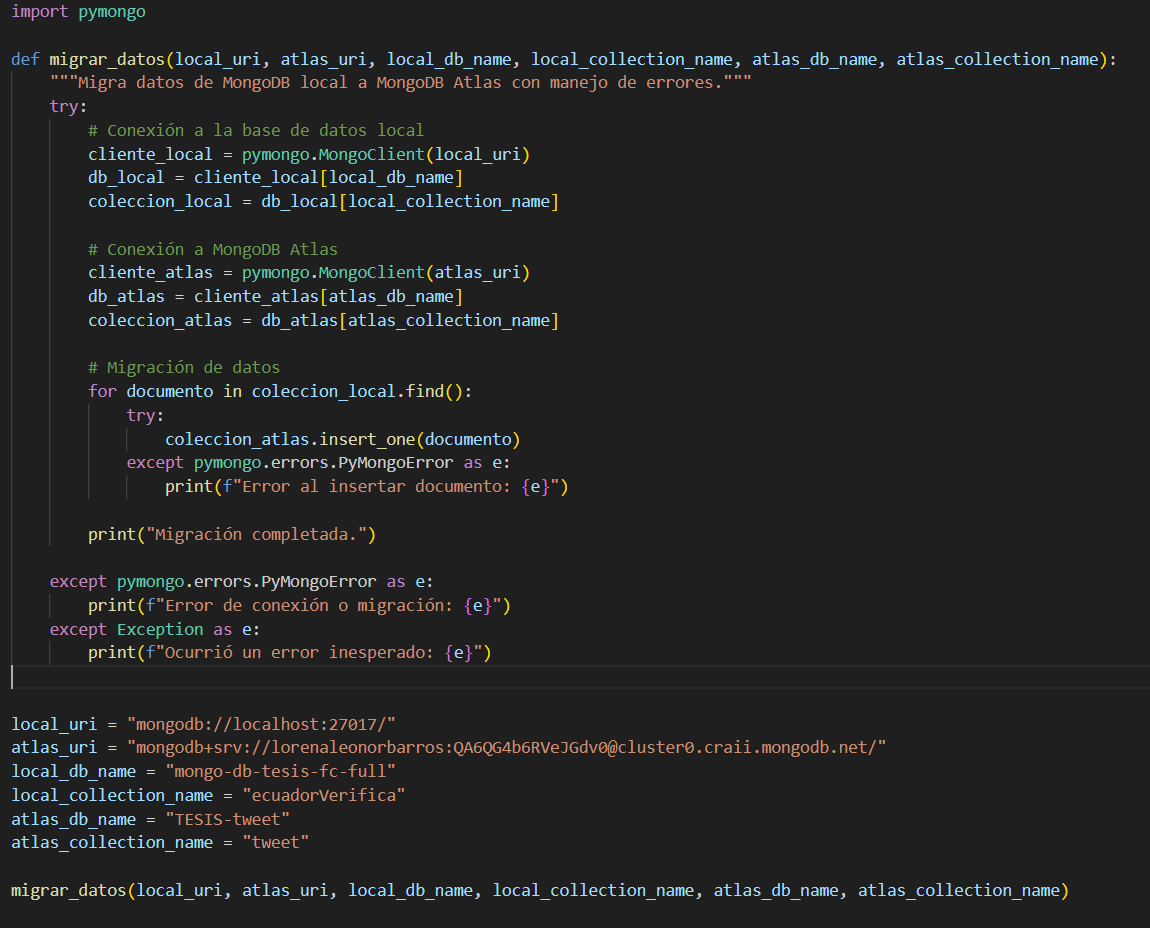
 En la etapa de carga de los datos en Mongo Atlas, primero se realiza la conexión en Python con la base local de MongoDB. Para cargar la carpeta que contiene los archivos json generados en el script de web Scraping.

**Figura 44.** Carga de archivos json en base local

Texto

El contenido generado por IA puede ser incorrecto.

**Figura 45.** Ejecutar la carga

Luego de eso con el script de Migración, se transfieren lo datos de la base de datos local a la base que esta alojada en la nube. Se realiza la conexión con la base de datos local, además se especifica la colección que se quiere migrar a la base de datos Mongo Atlas. También se realiza la conexión con la base de datos en la nube y se especifica a que colección se debe cargar los datos.

**Figura 46.** Migración de base de datos local a base alojada a nube

### Triggers para limpieza y transformación de datos

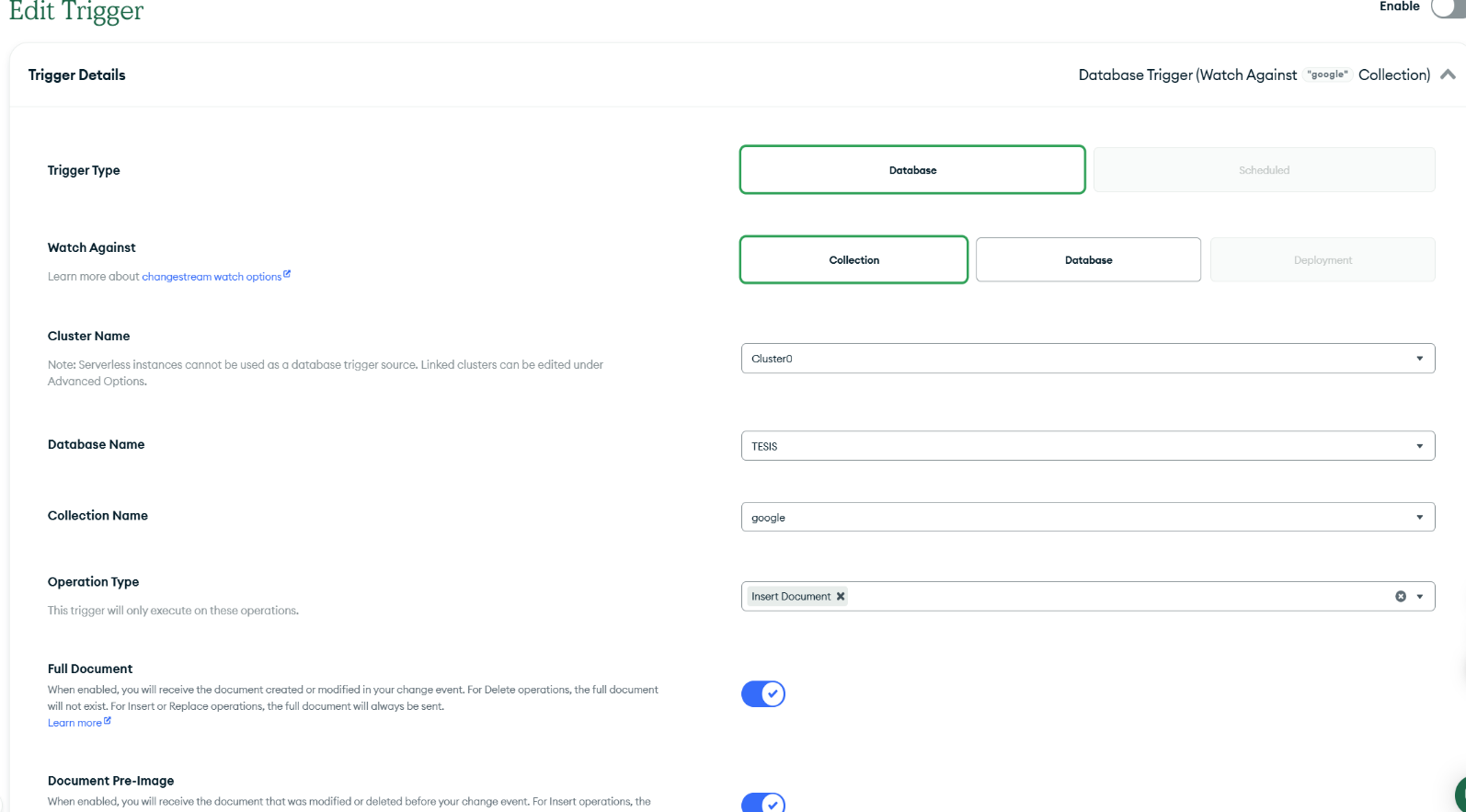
Para realizar la limpieza y transformación de datos se tienen dos disparadores para cada una de las fuentes de Google Fact Check Tools, corpus y de la red Social X.

Interfaz de usuario gráfica, Texto, Aplicación, Chat o mensaje de texto

El contenido generado por IA puede ser incorrecto.Para configurar los triggers también conocidos como disparadores o job, debemos de seleccionar triggers:

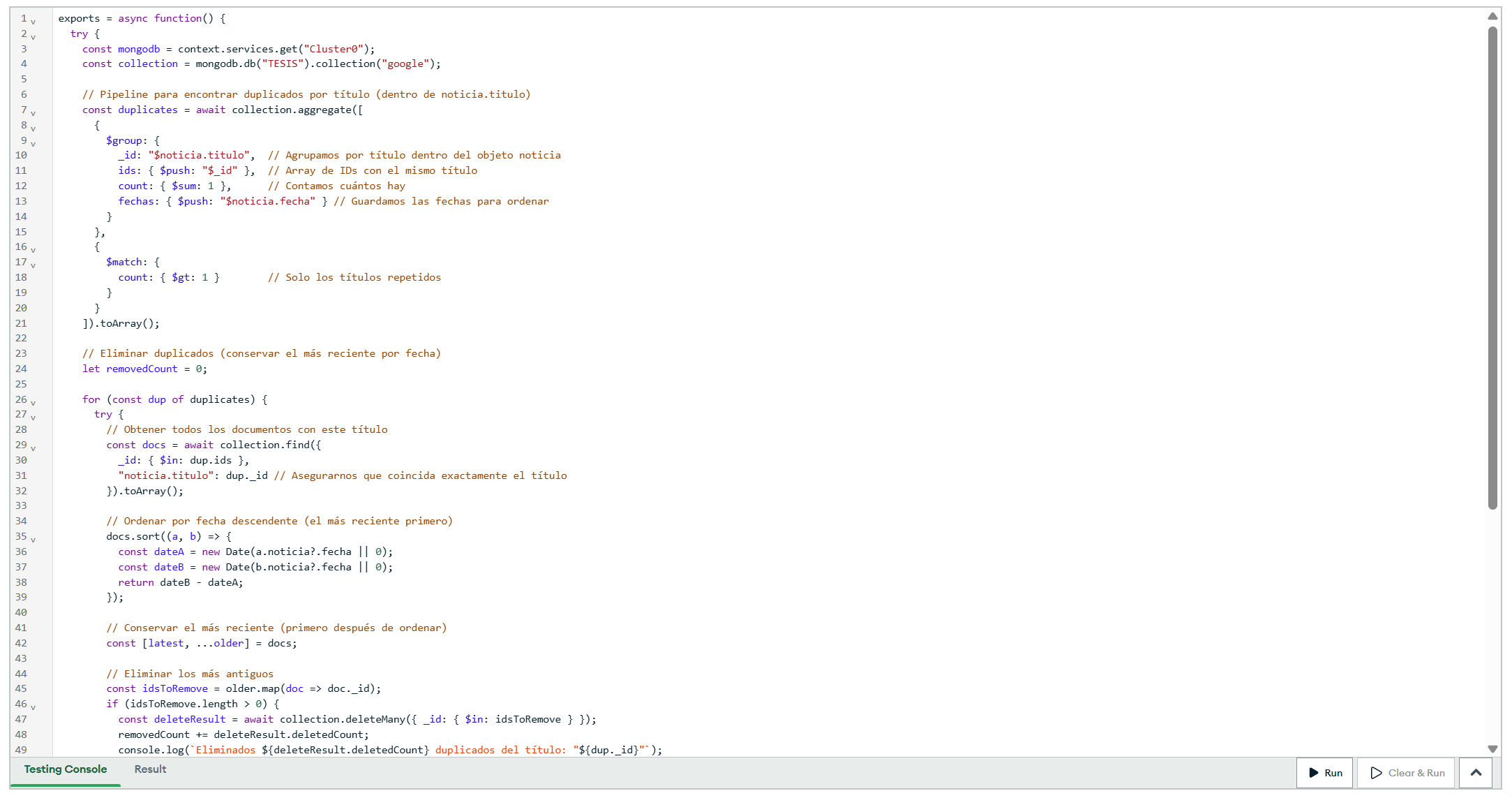
**Figura 47.** Disparadores

#### Trigger de limpieza para datos de Google Fact Check Tools

Luego debemos de seleccionar add Trigger y se abre la interfaz de donde se debe realizar la configuración. Para Google Fact Check Tools se realiza los siguientes Triggers, se tiene dos triggers uno de ellos se llama Limpieza Google y el otro Google\_transformacion. A continuación, se muestra la configuración del trigger de Limpieza Google:

**Figura 48.** Configuración de triggers Limpieza Google. Capturada realizada desde Mongo Atlas.

**Función**

****A continuación, se muestra la función que se ha colocado para realizar la limpieza de datos:

**Figura 49.** Función configurada en la limpieza de datos.

#### Trigger para la transformación de Google Fact Check Tools

Interfaz de usuario gráfica, Aplicación, Correo electrónico

El contenido generado por IA puede ser incorrecto.Para la transformación se utiliza un trigger llamado Google transformación, este trigger se diferencia del otro ya que es un disparador programado, se ejecuta en la fecha programada. El trigger Limpieza Google se ejecuta cada minuto.

**Figura 50.** Configuración del trigger para la transformación Google Transformacion

**Interfaz de usuario gráfica, Texto, Aplicación

El contenido generado por IA puede ser incorrecto.Función**

**Figura 51.** Función programada en Google Transformacion

#### Trigger para la limpieza de datos del Corpus

Para los datos del corpus también se realiza dos triggers uno de limpieza y otro de transformación.

Interfaz de usuario gráfica, Aplicación

El contenido generado por IA puede ser incorrecto.En este caso tenemos dos triggers programados para que se ejecuten en diferentes horarios. A continuación, se mostrará la configuración de limpieza de corpus, este trigger se le ha dado el nombre de “duplicadoscorpus”.

**Figura 52.** Configuración de trigger duplicadoscorpus.

Imagen que contiene Texto

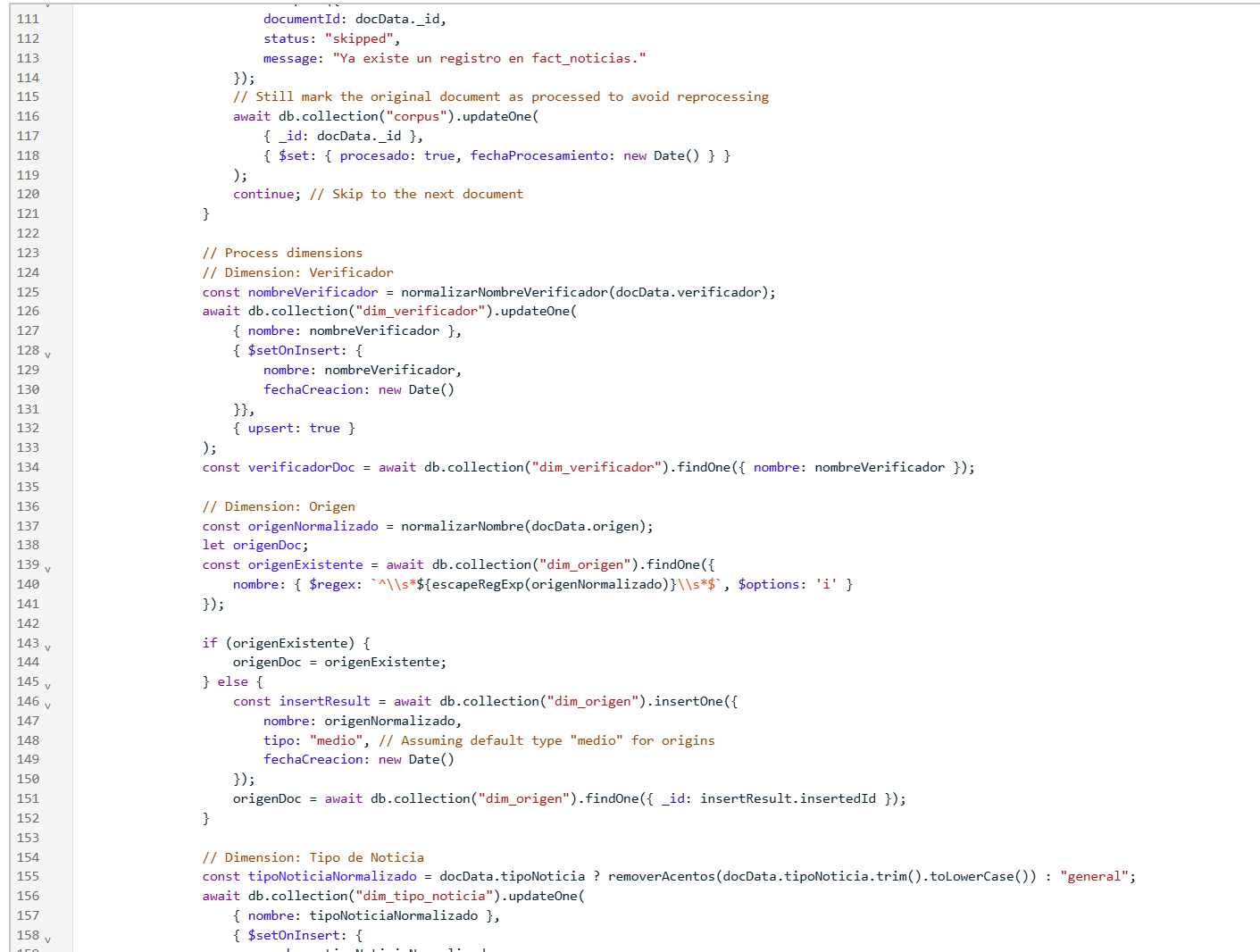
El contenido generado por IA puede ser incorrecto.A continuación, se muestra la función que se ha codificado:

**Figura 53.** Función de trigger duplicadoscorpus

#### Trigger para Transformación de datos del Corpus

Interfaz de usuario gráfica, Aplicación

El contenido generado por IA puede ser incorrecto.Para la transformación se realizó un trigger programado llama CORPUS, este trigger realiza las tablas de dimensiones y de hechos, ya que como los datos que recopilamos inicialmente están crudos.

**Figura 54.** Configuración de la hora a ejecutarse

**Figura 55.** Función realizada en el trigger Corpus

#### Trigger para la limpieza de datos de la Red Social X

Para la data recopilada de la red Social X, se tiene dos triggers uno que se ejecuta cada minuto este se le ha dado el nombre como triggers\_no\_verificados\_duplicados. El otro trigger llamado Trigger\_programado\_X este se lo ha configurado para ejecutar una vez a la semana a la hora especificada.

Interfaz de usuario gráfica, Aplicación

El contenido generado por IA puede ser incorrecto.La configuración que se muestra a continuación es sobre el **triggers\_no\_verificados\_duplicados**

**Figura 56.** Configuración triggers\_no\_verificados\_duplicados

Texto

El contenido generado por IA puede ser incorrecto.

**Figura 57.** Función en trigger ejecutado cada minuto triggers\_no\_verificados\_duplicados

#### Trigger para la tranformacion de datos de la red Social X

Para el trigger programado llamado Trigger\_programado\_x este realiza la transformación, creación de tablas de dimensiones y hechos

Interfaz de usuario gráfica

El contenido generado por IA puede ser incorrecto.

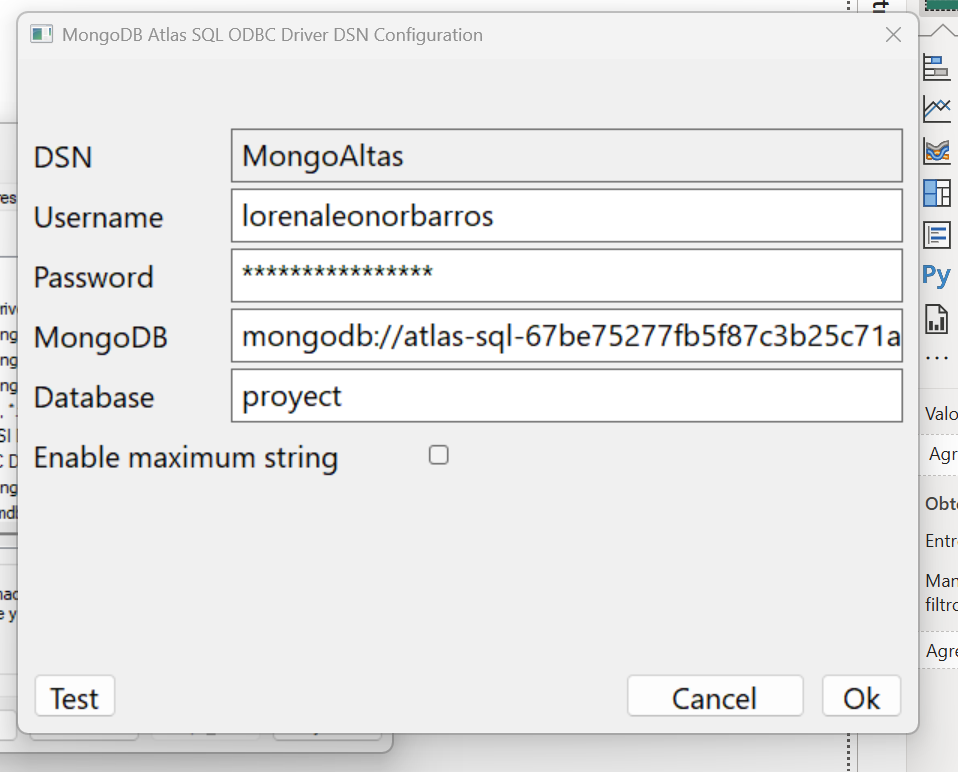
Imagen que contiene Texto

El contenido generado por IA puede ser incorrecto.**Figura 58.** Trigger programado X

**Figura 59.** Función programado en el trigger

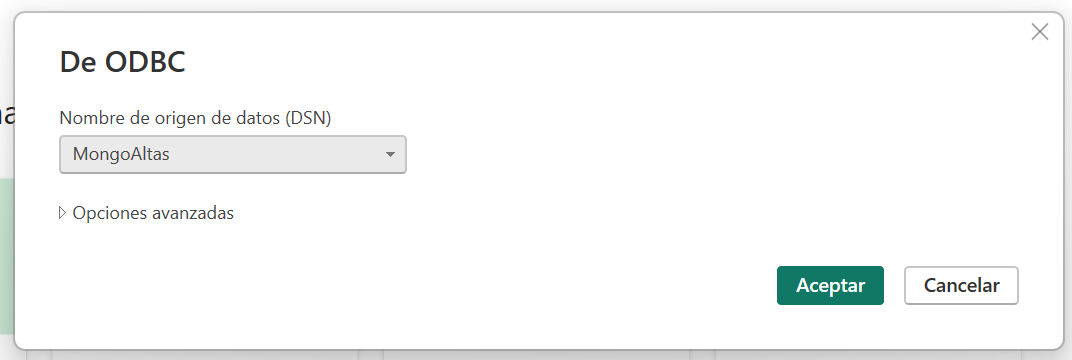
### Conexión de los datos con Power BI

Luego de realizar la limpieza y transformación de datos, se procede a realizar la conexión de los datos en Power BI se utiliza el ODBC. Antes de esto se debe realizar una configuración de este ODBC, se coloca el username, password y la uri y el database



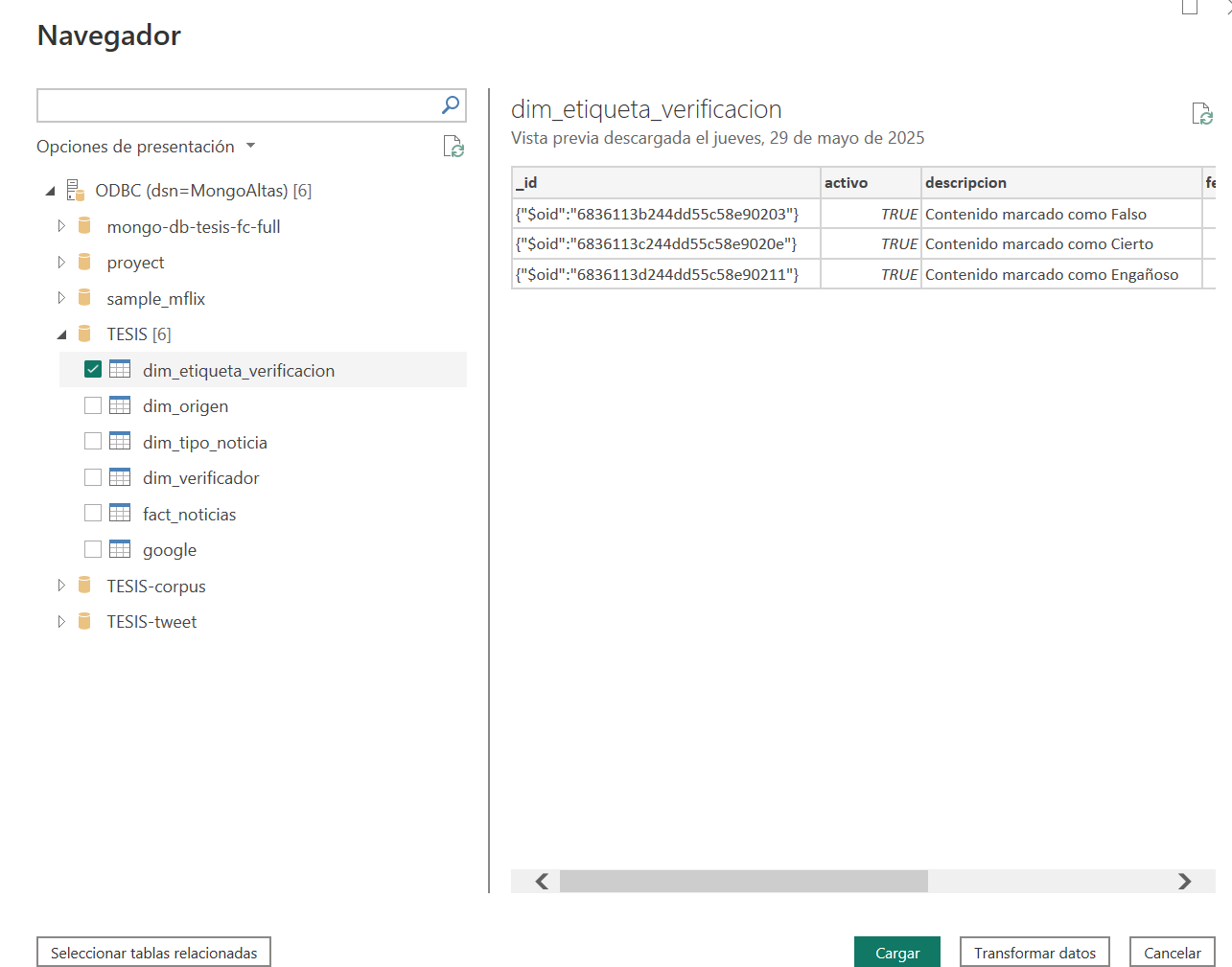
**Figura 60.** Configuración de odbc

Luego de esa configuración se ingresa a Power BI, y aparece el odbc que configuramos damos en aceptar



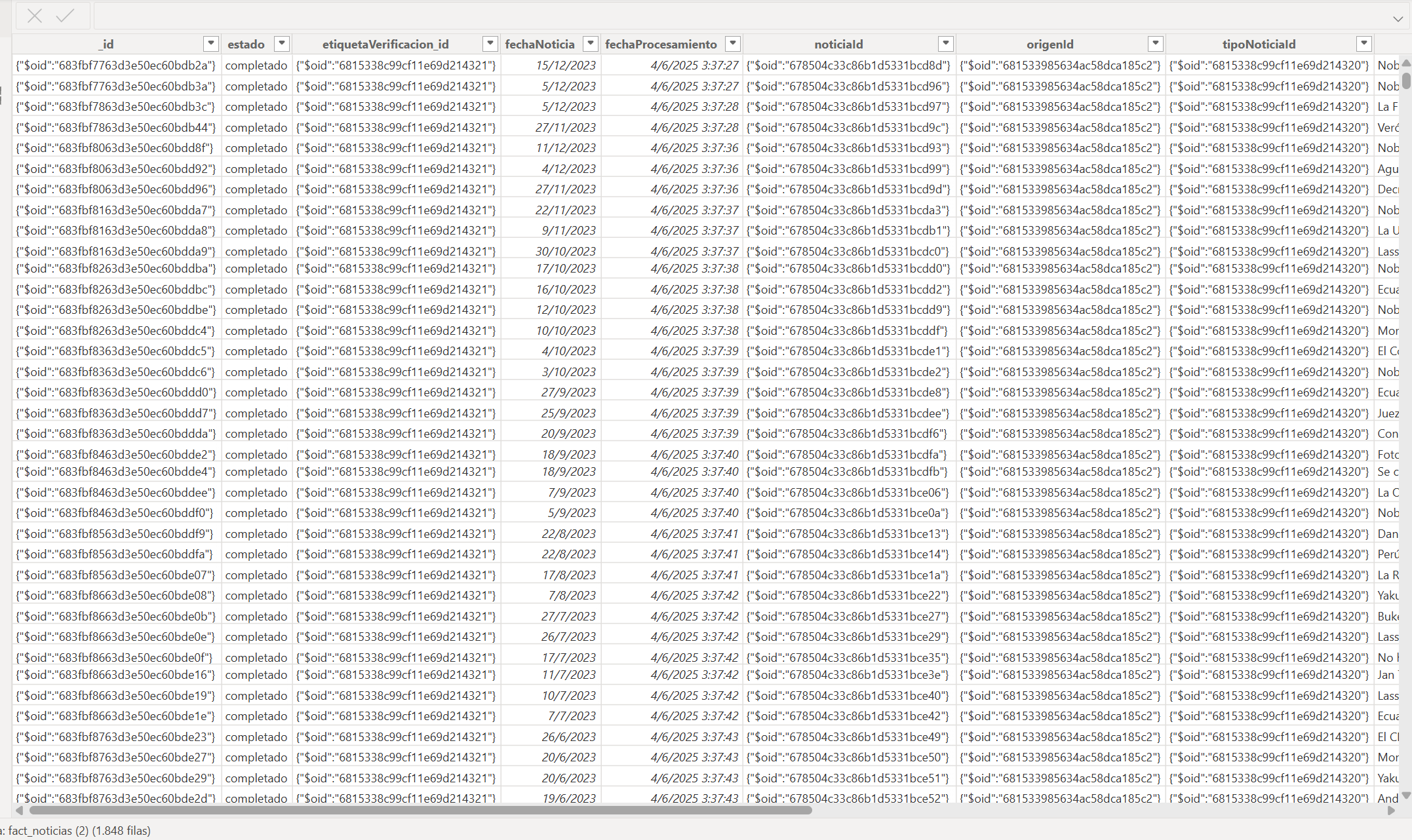
**Figura 61.** Origen de datos ODBC en Power BI

Después de dar aceptar, nos aparecerá las colecciones que se tiene en el cluster de MongoDB Atlas, se podrá seleccionar las colecciones a utilizar.



**Figura 62.** Colecciones de Mongo Atlas

Una vez cargada las colecciones que servirá para la creación de los dashboard, se procede a realizar los reportes en Power BI



**Figura 63.** Datos cargados

## Análisis preliminar del proyecto

Por medio del análisis preliminar del proyecto se examinaron características como la evaluación técnica, rentabilidad financiera y la capacidad operativa.

### Evaluación técnica

En este apartado incluirá la evaluación de las tecnologías, en el momento en el que se realiza el scraping de las fuentes de información verificada.

#### Obtención de los datos

* **Web scraping a través de Python:** Se empleo librerías selenium, ChromeDriver, pytesseract, json para extraer los datos de Google Fact Check Tools y de la red Social X.

#### Entorno tecnológico

* **Servidores en la nube para la base de datos NoSQL de Mongo Atlas:** La utilización del cluster para almacenar la data recopilada y programar triggers para realizar la limpieza y transformación, y así poder conectarla con Power BI para realizar el dashnoard.

#### Importación de la data

Por medio de Power BI se podrá importar esta data desde Mongo Atlas.

#### Capacidad de Modelado y visualización

Al momento de tener todos los datos listos se procede a realizar el modelado de datos para ver cómo están relacionadas las tablas y así tener la capacidad de crear visualizaciones**.**

#### Power BI Service

* Una vez culminado el dashboard se publicará para que las demás personas puedan visualizar

#### Hardware para desarrolladores

**Tabla 42.** Hardware utilizado para los desarrolladores

|  |  |
| --- | --- |
| **Hardware utilizado por los desarrolladores** | |
| **Nombre de equipo** | **Características** |
| Laptop Dell Inspiron 15 | **Procesador:** 12th Gen Intel(R) Core(TM) i7-1255U, 1700 Mhz. |
| **Disco duro:** 954 GB |
| **RAM:** Memoria física instalada (RAM) 16,0 GB |
| **SO:** Microsoft Windows 11 Home |
| Laptop Dell Inspiron 16 | **Procesador:** Procesador 13th Gen Intel(R) Core(TM) i7-1360P, 2200 Mhz, 12 procesadores principales, 16 procesadores lógicos |
| **Disco duro:** 931 GB |
| **RAM:** Memoria física instalada (RAM) 16,0 GB |
| **SO:** Microsoft Windows 11 Home |

Información realizada por los autores.

### Rentabilidad financiera

Proporciona la implementación mínima simplificando recursos.

### Capacidad operativa

Se comprobará que la generación del dashboard elaborado en Power BI contenga las especificaciones funcionales y no funcionales.

**Tabla 43.** Gastos salariales

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Gastos salariales** | | | | |
| **Rol** | **Cantidad** | **Plazo en meses** | **Honorario** | **Total** |
| **Lider de proyecto** | 1 | 3 | 0 | 0 |
| **Analista programador** | 2 | 3 | 0 | 0 |

Información realizada por los autores.

### Recursos

#### Recursos físicos

**Componentes físicos**

* **Servidor:** Dispositivo para mantener alojada la base de datos NoSQL (Mongo Atlas) esta base se encuentra alojada en la nube
* **Hardware:** 2 Laptops, herramientas para la implementación de Power BI y Power BI Service**.**

**Software**

* **Herramientas para el desarrollo del web Scraping:** Python, librerías como selenium, pytesseract entre otras.
* **Herramientas para realizar visualizaciones:** Power BI se utiliza la licencia gratuita para este proyecto
* **Cloud Services:** Se utiliza Mongo Atlas por su facilidad de manipular. También se utiliza el Servicio de Power BI para publicar el dashboard en la nube.
* **Navegador:** Se hace uso del navegador Google Chrome para el proceso automatizado del scraping.

#### Recursos Humano

**Equipo de desarrollo:**

* **Desarrolladores del web Scraping:** Son aquellos responsables de realizar la codificación para extraer los datos de las diferentes fuentes.
* **Analistas de datos:** Son aquellos profesionales que estarán encargados de extraer los datos, realizar la limpieza y también realizar el análisis de los datos.

### Presupuesto y financiamiento del proyecto

Se presenta a continuación los recursos que se utilizarán para el desarrollo del proyecto, con su respectivo presupuesto.

**Tabla 44.** Infraestructura para el desarrollo del proyecto

|  |  |  |
| --- | --- | --- |
| Infraestructura por utilizar | Uso | Costo |
| Laptop Dell Inspiron 16 5630 13th Gen Intel(R) Core (TM) i7-1360P 2.20 GHz | Se utilizará para el desarrollo del Dashboard | $0 |
| Laptop Dell 12th Gen Intel(R) Core (TM) i7-1255U, 1700 Mhz | Se utilizará para el web Scraping | $0 |
| Acceso a internet | Se utiliza para presentar avance, para manipular datos ya que están alojados en la nube y para compilar el web Scraping ya que estos datos se obtienen de la red Social X y Google Fact Check Tools | $0 |
| Power BI Service | Se utilizará para alojar el dashboard en la nube | $0 |

Nota. Información realizada por los autores.

## Fases de Diseño

### Modelo de Dominio

**Diagrama

El contenido generado por IA puede ser incorrecto.**

**Figura 64.** Modelo de Dominio para la red Social X. Realizado por los autores.

**Diagrama

El contenido generado por IA puede ser incorrecto.**

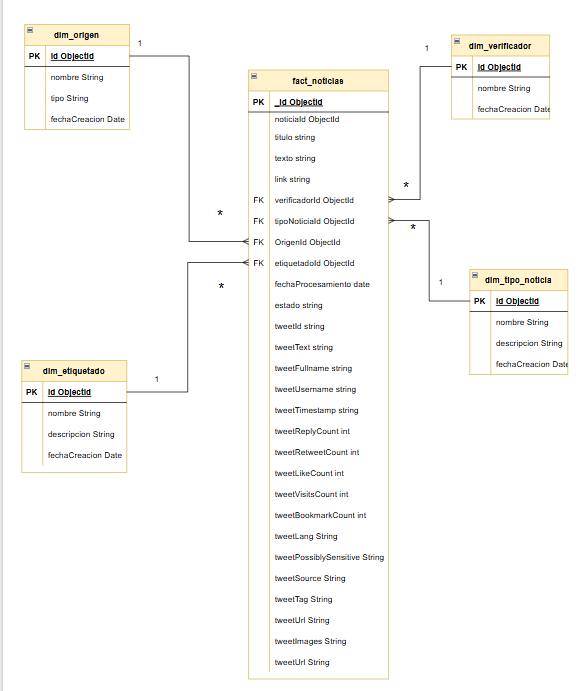
**Figura 65.** Modelo de Dominio del Corpus UG. Realizado por los autores.

**Diagrama

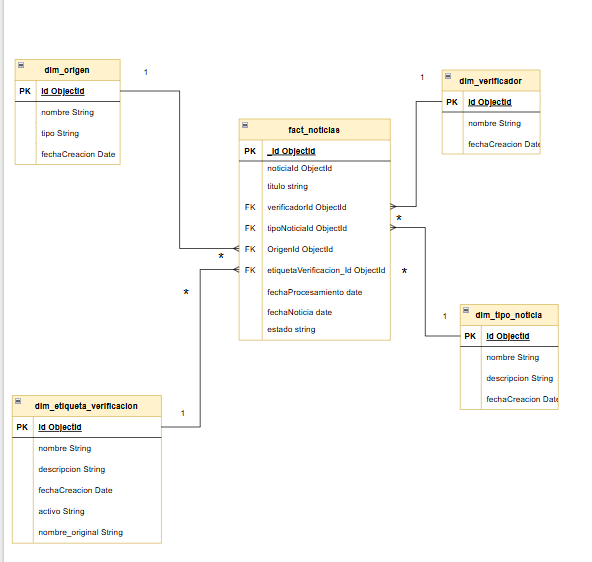
El contenido generado por IA puede ser incorrecto.**

**Figura 66.** Modelo de Dominio para las noticias obtenidas de Google Fact Check Tools.

### Diagrama de bases de datos



**Figura 67.**Diagrama de las colecciones de la red Social X. Elaborado por los autores

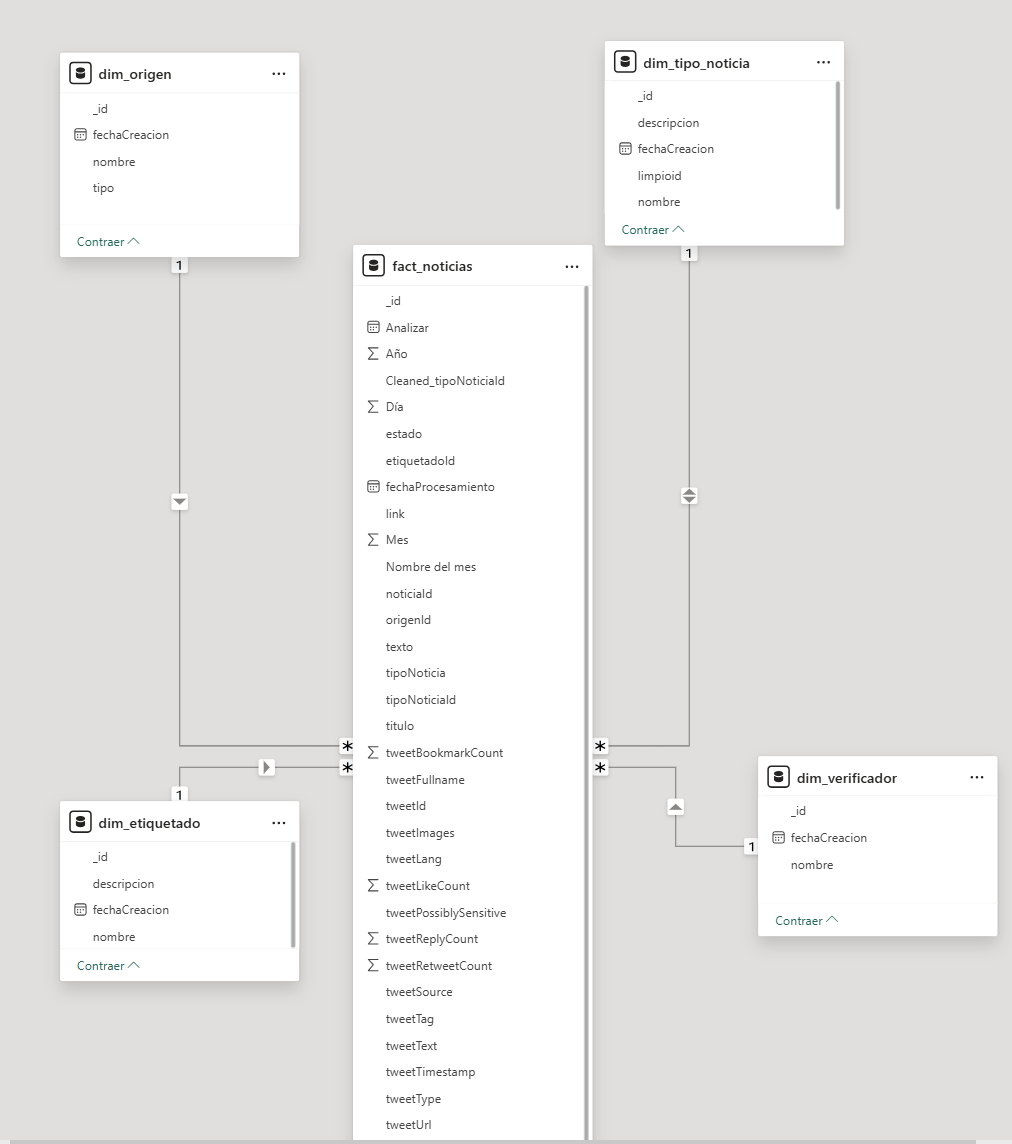


**Figura 68**. Diagrama de las colecciones del Corpus. Elaborado por los autores.

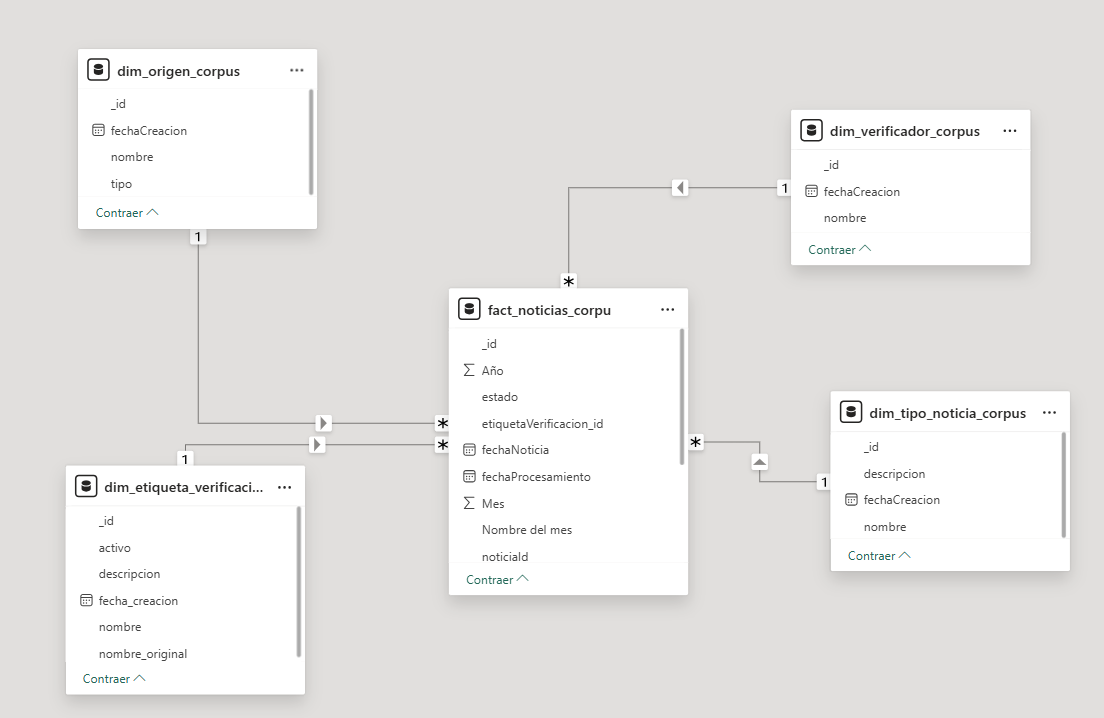


**Figura 69.** Diagrama de las colecciones de Google Fact Check Tools. Elaborado por los autores.

### Modelamiento de datos



***Figura 70.*** *Modelamiento de datos de la red Social X. Realizado por los autores.*



**Figura 71.** Modelamiento de datos del Corpus. Realizado por los autores.

**BIBLIOGRAFÍA**

Asamblea Nacional. (2023). *Ley Orgánica de Transparencia y acceso a la información pública* . https://www.dpe.gob.ec/wp-content/dpedocumentoslotaip/lotaip2023/ro-lotaip-2023.pdf

Ashraf Alameer. (2019). *Data visualization using Power BI and MongoDB*. https://medium.com/@alameerashraf/data-visualization-using-power-bi-and-mongodb-e694328220b5

Calero Katherine. (2021). *¿Noticia falsa? ¿Qué es eso?* https://uees.edu.ec/noticia-falsa-que-es-eso/

Casanoves, J. L. (2021). *La inteligencia de negocios como una oportunidad clave para las empresas*. https://riunet.upv.es/bitstream/handle/10251/174827/Legido%20-%20La%20inteligencia%20de%20negocios%20como%20una%20oportunidad%20clave%20para%20las%20empresas.pdf?sequence=1

Castillo-Riquelme, V., Hermosilla-Urrea, P., Poblete-Tiznado, J. P., & Durán-Anabalón, C. (2021). Fake news and unfunded beliefs in the post-truth age. *Universitas*, *34*, 87–108. https://doi.org/10.17163/UNI.N34.2021.04

César Bartolomé, M. (2021). Redes sociales, desinformación, cibersoberanía y vigilancia digital: una visión desde la ciberseguridad. *RESI: Revista de Estudios En Seguridad Internacional, ISSN-e 2444-6157, Vol. 7, No. 2, 2021, Págs. 167-185*, *7*(2), 167–185. https://doi.org/10.18847/1.14.9

Cruz, C., & Muñiz, J. (2022). La integración de datamart con datawarehouse. *UNESUM-Ciencias. Revista Científica Multidisciplinaria.*, *6*(1), 23–30. https://doi.org/10.47230/unesum-ciencias.v5.n4.2021.470

Curto, J. (2020). *Introducción al business intelligence y big data*. https://openaccess.uoc.edu/bitstream/10609/150184/3/IntroduccionAlBusinessIntelligenceYBigData.pdf

Digari Dharmendra. (2024). *What is Power Query in Power BI? - A Guide for Beginners*. https://www.whizlabs.com/blog/power-query-power-bi/

Ecuador Chequea. (2022). *Metodologia - Ecuador Chequea*. https://ecuadorchequea.com/metodologia/

Ecuador Verifica. (2020). *METODOLOGÍA - Ecuador Verifica*. https://ecuadorverifica.org/metodologia/

Esther Paniagua. (2021). *Big data El poder de los datos*. www.fundacionbankinter.org

Ferrer, J. V. (2020, October 7). *Como crear tu propio cuadro de mando en Power BI Desktop*. https://inforges.es/blog/webinar-gratuito-como-crear-tu-propio-cuadro-de-mando-en-power-bi-desktop/

García-Jiménez, A. de J., Aguilar-Morales, N., Hernández-Triano, L., & Lancaster-Díaz, E. (2021). La inteligencia de negocios: herramienta clave para el uso de la información y la toma de decisiones empresariales. *Revista de Investigaciones Universidad Del Quindío*, *33(1)*, 132–139. https://revistas.uniquindio.edu.co/ojs/index.php/riuq/article/view/514/479

Goel, P. (2023). Scalable ETL Processes in Azure Data Factory: Best Practices for Data Engineers. *Researchgate.Net*, *11*. https://www.researchgate.net/profile/Swathi-Garudasu/publication/389435369\_Scalable\_ETL\_Processes\_in\_Azure\_Data\_Factory\_Best\_Practices\_for\_Data\_Engineers/links/67c21b70f5cb8f70d5c3dd3a/Scalable-ETL-Processes-in-Azure-Data-Factory-Best-Practices-for-Data-Engineers.pdf

Google. (n.d.). *About Fact Check Tools*. Retrieved March 17, 2025, from https://toolbox.google.com/factcheck/about

Grupo Goberna. (2025). *Granjas de Trolls en las Elecciones Presidenciales de Ecuador 2025: Desinformación y Manipulación Digital*. https://grupogoberna.com/granjas-de-trolls-elecciones-ecuador-2025/

Hernandez Mendoza, S. L., & Duana Avila, D. (2020). Técnicas e instrumentos de recolección de datos. *Boletín Científico de Las Ciencias Económico Administrativas Del ICEA*, *9*(17), 51–53. https://doi.org/10.29057/ICEA.V9I17.6019

International Fact-Checking Network. (2025). *Signatories*. https://ifcncodeofprinciples.poynter.org/signatories

José M. García-Miguel. (2022). Lingüística de corpus: de los datos textuales a la teoría lingüística. *Estudios de Lingüística Del Español,* *45*, 11–42. https://www.raco.cat/index.php/Elies/article/download/403735/497550

Leo, H., & James, A. (2024). *A Review of Open-Source ETL (Extract, Transform, Load) Tools for Data Integration.* https://www.researchgate.net/publication/389356972\_A\_Review\_of\_Open-Source\_ETL\_Extract\_Transform\_Load\_Tools\_for\_Data\_Integration/citation/download

Ley Orgánica de Comunicación. (2019). *Ley Orgánica de Comunicación*. https://www.telecomunicaciones.gob.ec/wp-content/uploads/2020/01/Ley-Organica-de-Comunicaci%C3%B3n.pdf

López Espinoza, A. I. (2024). *Nuevo modelo de datos e implementación de subflujos en proceso ETL para aumentar la eficiencia operativa en NeoSoft* [Universidad de Chile]. https://repositorio.uchile.cl/handle/2250/202814

Marrero, L. (2023). *Un estudio comparativo de bases de datos NoSQL (Doctoral dissertation, Universidad Nacional de La Plata).*

Microsoft. (2024). *Tipos de visualización en Power BI*. https://learn.microsoft.com/es-es/power-bi/visuals/power-bi-visualization-types-for-reports-and-q-and-a

Microsoft. (2025, July 2). *Puerta de enlace de datos local - Power BI*. https://learn.microsoft.com/en-us/power-bi/connect-data/service-gateway-onprem

Microsoft Learn. (2024a, October 14). *Usar DirectQuery en Power BI Desktop - Power BI*. https://learn.microsoft.com/es-es/power-bi/connect-data/desktop-use-directquery

Microsoft Learn. (2024b, December 30). *Ajuste de tamaño de la puerta de enlace de datos local - Power BI | Microsoft Learn*. https://learn.microsoft.com/es-es/power-bi/guidance/gateway-onprem-sizing

Microsoft Learn. (2025, February 5). *¿Qué es Power Query?* https://learn.microsoft.com/es-es/power-query/power-query-what-is-power-query

Newman, N., Fletcher, R., Robertson, C. T., Eddy, K., & Kleis Nielsen, R. (2022). *Reuters Institute Digital News Report 2022*. https://doi.org/10.60625/risj-x1gn-m549

Niño Montero José Segundo, & Mendoza Hidalgo Mary Liz. (2021). *La investigación científica en el contexto académico* (Infinite Study, Ed.). NSIA Publishing House Editions. https://www.google.com.ec/books/edition/La\_investigaci%C3%B3n\_cient%C3%ADfica\_en\_el\_cont/B7koEAAAQBAJ?hl=es&gbpv=1

Oracle. (2025). *¿Qué es un almacén de datos?* https://www.oracle.com/cr/database/what-is-a-data-warehouse/

Ortega, J. (2025). *Ingeniería de datos.  Diseño, implementación y optimización de flujos de datos en Python*. Ra-ma S.A. Editorial y Publicaciones.

Palacios, I., & Cusot, G. (2019). Las fake news y las estrategias de verificación del discurso público: Caso Ecuador Chequea. *#PerDebate*, *3(1)*, 88–107. https://doi.org/10.18272/PD.V3I1.1558

Pineda Solorio, M. E. (2024). El impacto de las fake news en las redes sociales. *Revista Electrónica Sobre Tecnología, Educación y Sociedad*, *11*. https://orcid.org/0000-0002-1069-8946

Poynter. (2025). *Red Internacional de Verificación de Datos - Poynter*. https://www.poynter.org/ifcn/

Regina de Miguel. (2023). *Comparativa: 11 plataformas de business intelligence (BI)*. https://revistabyte.es/comparativa/plataformas-de-business-intelligence-bi/

Reyes Mena, F. X. (2023). *Generación de data warehouse para implementación de business intelligence que permita visualizar el comportamiento de los clientes y tomar medidas de acción comercial*. https://repositorio.puce.edu.ec/server/api/core/bitstreams/dbf97d80-3341-4c3e-9b56-ff8281053793/content

Rodríguez, M. D. Pérez. (2024a). *Business Intelligence* (ICB Editores, Ed.). [s.n.].

Rodríguez, M. D. Pérez. (2024b). *Business Intelligence María Dolores Pérez Rodríguez* (ICB Editores, Ed.). https://www.google.com.ec/books/edition/Business\_Intelligence/ChQ8EQAAQBAJ?hl=es&gbpv=0

Sánchez Duarte, J. M., & Magallón-Rosa, R. (2023). Desinformación. *Revista En Cultura de La Legalidad*, *24*, 236–249. https://e-revistas.uc3m.es/index.php/EUNOM/article/download/7663/5986/

Sanchez Lennini. (2024). *¿Qué es Business Intelligence? - Initium*. https://www.initiumsoft.com/blog\_initium/business-intelligence/

Sarango, A., Baldeón, M., Medina, M., Gavilanes, E., & Burbano, M. (2025). *Inteligencia de Negocios: Principios Fundamentales y Aplicaciones Empresariales: Business Intelligence: Fundamental Principles and Business Applications*. Know Press. https://doi.org/https://doi.org/10.70180/978-9942-7273-8-1

Sign Consulting. (2023, October 11). *Las 7 Principales Plataformas de Business Intelligence: Una Perspectiva de Sign Consulting - LinkedIn*. https://www.linkedin.com/pulse/las-7-principales-plataformas-de-business-intelligence/

Skender, F., & Manevska, V. (2022). Data Visualization Tools-Preview and Comparison. *Journal of Emerging Computer Technologies*, *2*(1), 30–35. https://dergipark.org.tr/en/download/article-file/2503669

Slusarczyk Antosz, M. (2024). *Inteligencia de Negocios -Power bi* (La Caracola Editores, Ed.; Vol. 17). Decanato de Publicaciones. http://cimogsys.espoch.edu.ec/direccion-publicaciones/public/docs/books/2024-04-30-163812-Inteligencia\_de\_negocios.pdf

Tarale, M. H. (2025). Web Scraping using Python. *The Academic*, *3*, 1023–1024. https://doi.org/https://doi.org/10.5281/zenodo.14852767

Toapanta Bernabe, M., Garcia-Cumbreras, M. A., & Urena-Lopez, L. A. (2024). Uso de modelos en idioma español para la detección de noticias falsas y verificación de hechos en tuits de Ecuador Chequea y Ecuador Verifica. *Revista Tecnológica - ESPOL*, *36*(2), 158–173. https://doi.org/10.37815/rte.v36n2.1219

Ulchur Iván. (2024). *Ecuador, entre la violencia y la desinformación - The New York Times*. https://www.nytimes.com/es/2024/01/28/espanol/opinion/ecuador-violencia-desinformacion.html

Vélez Bermello, G. (2024). Fact-checking: campaña política de elecciones seccionales en Ecuador. Caso Ecuador Verifica. *ComHumanitas: Revista Científica de Comunicación*, *15*(2), 214–232. https://doi.org/10.31207/RCH.V15I2.458

Vélez Bermello, G., & Henríquez Coronel, P. (2024). Fact-checking: concepto sobre verdad y las prácticas en Ecuador chequea y Ecuador verifica. *Hologramática*, *40(3)*, 3–29. http://revistas.unlz.edu.ar/ojs/index.php/rholo/article/view/124

Vélez Bermello, G. L. (2020). Inmediatez y fact-checking: análisis del portal Ecuador Chequea. *Revista ABRA*, *40*(61), 63–87. https://doi.org/10.15359/abra.40-61.3

X Corp. (2017). *Reglas de desarrollo de automatización* . https://help.x.com/es/rules-and-policies/x-automation

X Corp. (2025a). *Autenticidad*. https://help.x.com/es/rules-and-policies/authenticity

X Corp. (2025b). *Las reglas de X*. https://help.x.com/es/rules-and-policies/x-rules

Zhang, Q. (2024). The Impact of Interactive Data Visualization on Decision-Making in Business Intelligence. *Advances in Economics, Management and Political Sciences*, *87*(1), 166–171. https://doi.org/10.54254/2754-1169/87/20241056