



**UNIVERSIDAD
DE ANTIOQUIA**

**Implementación y construcción de vistas Front-end y base
de datos normalizada para la plataforma de marcado de
autogestión de riesgos -MAR**

Autor

Sebastián Buriticá Montoya

Universidad de Antioquia

Facultad de Ingeniería

Medellín, Colombia

2020



**Implementación y construcción de vistas Front-end y base de datos normalizada para la
plataforma de mercado de autogestión de riesgos (MAR)**

Sebastian Buritica Montoya

Práctica empresarial como requisito parcial para optar al título de:

Ingeniero de Telecomunicaciones

Asesor interno:

Natalia Gaviria Gómez

Universidad de Antioquia

Facultad de Ingeniería

Medellín, Colombia

2020

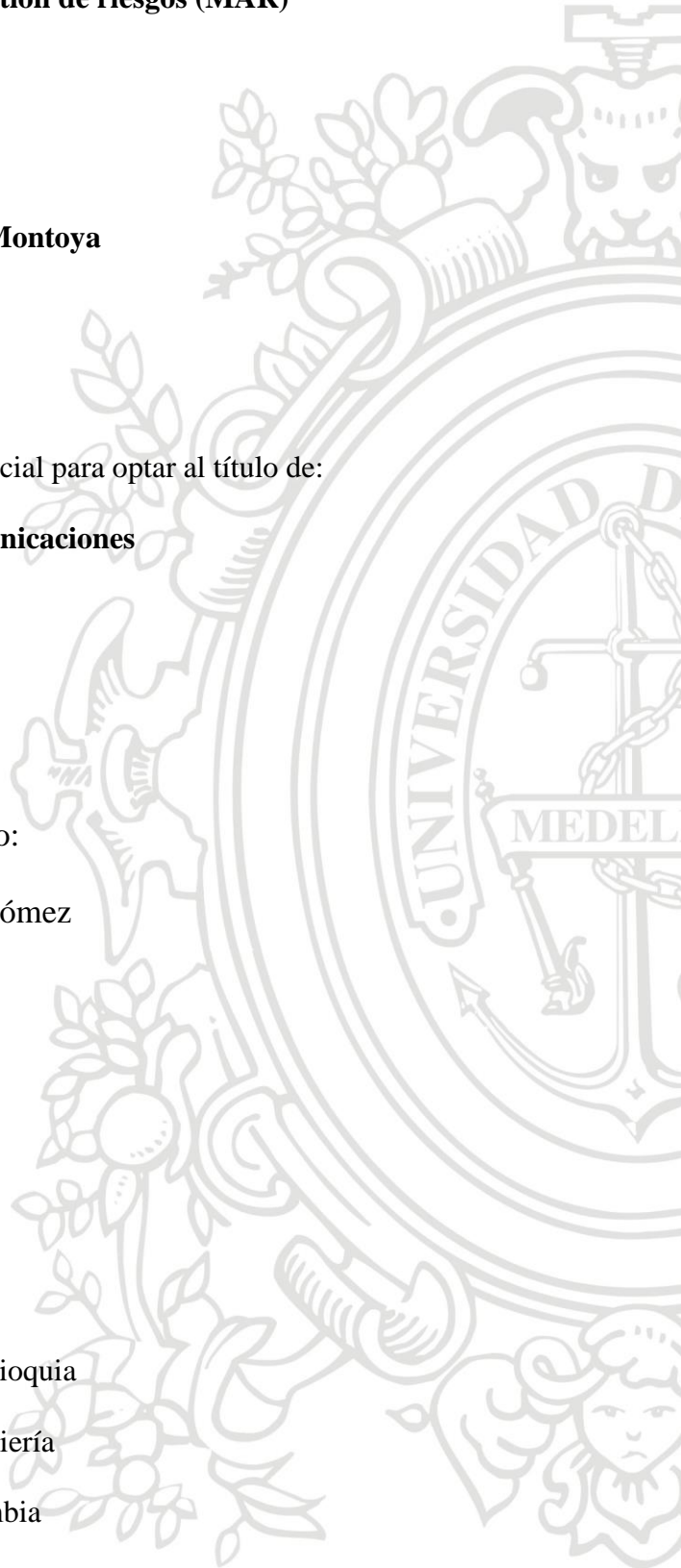


Tabla de contenido

Resumen.....	4
1. Introducción.....	5
2. Objetivos.....	7
2.1. Objetivo general.	7
2.2. Objetivos específicos.....	7
3. Marco teórico.....	8
3.1. Analítica de datos.	8
3.2. Soluciones RegTech.....	9
3.3. Front-end.	10
3.4. Lenguajes de programación HTML, JavaScript y CSS.	10
3.5. Base de datos.	11
3.6. Normalización de bases de datos.	13
3.7. Otros conceptos importantes.	14
4. Metodología.....	15
4.1. Metodología de trabajo.....	15
4.2. Proceso de desarrollo.	17
4.3. Etapas del proyecto.	17
4.4. Fases del estudio previo.	18
4.4.1. Diseños previos.....	18
4.4.2. Arquitectura del proyecto.....	28
4.5. Fase de producción.....	30
4.5.1. Front-end.	30
4.5.2. Base de datos.....	8
5. Análisis de resultados	18
6. Conclusiones.....	27
7. Referencias	29

Implementación y construcción de vistas front-end y de base de datos normalizada para la plataforma de Mercados de Autogestión de Riesgos -MAR

Resumen

La gerencia de Investigación, Desarrollo e Innovación (I+D+I) perteneciente a la vicepresidencia de Auditoría Interna del Grupo Bancolombia propuso el desarrollo de una plataforma de Mercado de Autogestión de Riesgos –MAR–, RegTech, con el objetivo de proporcionar a los clientes una herramienta funcional para la prevención de riesgos, así como para facilitar la adaptación a los cambios. Debido a ello, se realizó un soporte y acompañamiento en el desarrollo del proyecto orientado al desarrollo front-end de la plataforma, el diseño y la organización de la base de datos y, el ingreso de información a esta mediante la técnica de Web scraping en diferentes listas de control propuestas y definidas por la misma entidad bancaria.

Para este desarrollo, fue necesario consultar la documentación sobre Django y Bootstrap, sobre los frameworks propuestos para el desarrollo back-end y front-end para la gestión de la base de datos, y el desarrollo de la interfaz de usuario (UI), respectivamente; también se desarrollaron habilidades en la técnica Web scraping para la extracción de información de las páginas web. Adicional, a lo largo de la práctica se definió la metodología Scrum, implementando sprints semanales con lo cual fue posible cumplir con los objetivos propuestos.

Finalmente, a partir del análisis de resultados fue posible concluir que el desarrollo de este proyecto permitió construir una primera versión funcional de la plataforma, amigable con el usuario y funcional en el tema de consulta de información.

1. Introducción

Bancolombia es un grupo financiero multinacional colombiano que hace presencia en el continente americano a través de diferentes filiales y unidades de negocio que operan principalmente en Centro y Sur de América. Debido a los desafíos que le impone el mercado, surge la necesidad de que el banco esté a la vanguardia en el ámbito tecnológico para dar soluciones eficientes y eficaces a las necesidades no solo de los usuarios sino de la empresa misma a través de la creación y el desarrollo de nuevos aplicativos. Se considera que las soluciones tecnológicas representan una oportunidad de ofrecerles a los usuarios facilidad y rapidez en la realización de transacciones desde cualquier parte, al mismo tiempo que permite la consecución de mejores resultados en el desarrollo de las actividades empresariales. De igual forma, se resalta la necesidad de implementar herramientas para proveer seguridad y protección a la información de manera rigurosa.

En este marco de referencia, resulta importante la implementación de herramientas de análisis y predicción de datos para propiciar estrategias de mercado y aumentar los ingresos y la eficiencia operativa, entre otras ventajas. En este sentido, dentro de los nuevos desarrollos tecnológicos, la analítica de datos ha desempeñado en los últimos años un papel de gran relevancia, siendo este un aspecto que ha sido reconocido por la Vicepresidencia de Auditoría Interna de Bancolombia y razón por la cual esta dependencia ha buscado herramientas que ayuden a mitigar el riesgo de fraude, una amenaza frecuente y de alta probabilidad.

Al respecto, cabe resaltar que la Gerencia de Investigación, Desarrollo e Innovación (I+D+I), se encarga del desarrollo de todas estas herramientas apoyándose

en un equipo de auditores especializados en temas de tecnología con el fin de plantear soluciones que, además de trascender las funciones típicas de la auditoría, contribuyan a la prevención y disminución de los riesgos ya mencionados.

De acuerdo con lo anteriormente planteado, el objetivo del presente proyecto consiste en diseñar y desarrollar el front-end de la plataforma web denominada MAR (Mercado de Autogestión de Riesgos) en la que el cliente podrá encontrar toda la información relevante de la persona o empresa que desee consultar, entregándole un perfil de riesgo de negocio, con lo cual se busca contribuir a la mitigación del riesgo de fraude en una posible asociación futura. La plataforma MAR se concibe como un servicio de uso exclusivo para los clientes y proveedores de Bancolombia, para el cual se hará un cobro.

El perfil de riesgo entregado por MAR consta de tres niveles: alto, medio o bajo, y se fundamenta en la clase y cantidad de listas en las que se encuentre la persona o empresa consultada. Para la definición del perfil, se parte de una clasificación en cuatro categorías:

- **Reputación:** aparición de la persona o empresa en distintos medios de comunicación de alrededor de 30 diferentes países, donde cada noticia tendrá la posibilidad de ser calificada por medio de un sistema de estrellas por los usuarios de la plataforma con el fin de entregar una calificación ponderada de la noticia.
- **LAFT:** listas basadas en temas de lavado de activos y financiación del terrorismo. Algunos ejemplos son la lista Clinton y la lista Duque. El nivel de afectación a la persona o empresa dependerá de la cantidad de listas en las que aparezcan.

- **Crédito:** listas entregadas por las centrales de riesgo, cuyos resultados afectarán negativamente a la empresa o persona si esta se encuentra en varias listas.
- **Antecedentes judiciales:** listas públicas de diversas entidades del Estado, tanto nacionales como internacionales: fiscalía, FBI, entre otras.

El desarrollo del presente proyecto se enmarca en la iniciativa de MAR, y se basa, en gran medida, en el concepto de RegTech, es decir, en herramientas e innovaciones tecnológicas usadas para mejorar los parámetros normativos de las empresas, optimizar distintos procesos e impulsar la eficiencia empresarial, entre otros. El primer paso para la ejecución de esta iniciativa consiste en desarrollar el front-end de dicha MAR a partir de la creación y normalización de su base de datos, para luego extraer la información de las listas, integrar el código desarrollado al proyecto y finalmente, dar forma a la plataforma web MAR.

2. Objetivos

2.1. Objetivo general.

Implementar el front-end de la plataforma MAR (Mercado de Autogestión de Riesgos) a partir de la creación y normalización de la base de datos.

2.2. Objetivos específicos.

- Desarrollar las vistas front-end mediante los lenguajes de programación HTML 5, CSS y JavaScript, a partir de un *mock up* entregado por un diseñador web.
- Integrar el código desarrollado a la plataforma creada para el proyecto.

- Extraer la información de las listas mediante la técnica de Web scraping con la ayuda del entorno de desarrollo Spyder.
- Diseñar y construir la base de datos normalizada en la que estará alojada la información de cada una de las listas desarrolladas con la técnica de Web scraping con la ayuda del Framework Django y el aplicativo MySQL Workbench 8.0.

3. Marco teórico

En años recientes, los avances tecnológicos, particularmente las soluciones RegTech, han ayudado a la industria y a las autoridades a abordar mejor los riesgos (tradicionales y nuevos) para la estabilidad financiera, la protección del consumidor y la integridad del sistema financiero. De igual modo, esta planeación del riesgo y prevención del fraude se ejecutan a partir de distintas herramientas que incluyen software, Big Data, Blockchain y Machine Learning, entre otras, además de diversos lenguajes de programación.

A continuación, se describen algunos de los conceptos, nociones y tendencias que orientarán el desarrollo de este proyecto.

3.1. Analítica de datos.

De acuerdo con Delgado (2014), en desarrollo de prototipos virtualizados de computación de alto desempeño (HPC), la analítica de datos consiste en la inteligencia o seguimiento a determinada información que puede incluir rasgos o perfilamiento de usuarios, procesos de una empresa, tendencias del mercado u otro tipo de datos que sirven para propósitos muy específicos. Por lo demás, la analítica de datos puede dividirse en cuatro pasos: análisis, procesamiento, categorización y utilización. En este

punto se destaca que la tecnología ha permitido que grandes volúmenes de información puedan ser tratados, procesados y almacenados para distintos fines en todas las ramas de la economía y la industria, incluyendo el sector financiero, un ejemplo claro de esto es la Big Data.

Este concepto, a su vez, engloba la noción de base de datos, definida como “una colección de información organizada de forma que un programa de ordenador pueda seleccionar rápidamente los fragmentos de datos que necesite. Una base de datos es un sistema de archivos electrónico” (masadelante.com, 2019).

3.2. Soluciones RegTech.

Según Fernández (s.f.), las empresas RegTech hacen alusión a aquellas organizaciones o emprendimientos que utilizan diferentes soluciones tecnológicas con el fin de establecer y mejorar los parámetros normativos de estas. Precisamente, RegTech proviene del inglés *regulation* (regulación, norma) y *technology* (tecnología). Varias soluciones RegTech incluyen Big Data, Machine Learning, Blockchain, el llamado Internet de las cosas y la tecnología de reconocimiento en gestión de riesgos de los estatutos vigentes.

De manera general, RegTech proporciona soluciones tecnológicas en varias facetas del cumplimiento normativo. Al respecto, Weber (2017) menciona que:

- Gestionar requisitos reglamentarios y los imperativos de cumplimiento mediante la identificación de los impactos de las disposiciones reglamentarias en los modelos de negocio, productos y servicios, actividades funcionales, políticas, procedimientos operativos y controles.

- Permitir sistemas y datos empresariales compatibles.
- Controlar y gestionar riesgos regulatorios, financieros y no financieros.
- Realizar informes de cumplimiento normativo.
- Crear perfiles de riesgo LAFT (Lavado de Activos y Financiación de Terrorismo) para reducir falsos positivos y minimizar operatividad

3.3. Front-end.

Consiste en la parte de una página web que interactúa con los usuarios que la visitan. Es decir, es la parte visible, la que muestra los contenidos, el diseño y las herramientas que le permiten al usuario navegar por la página web como lo deseen. También puede ser definido como la parte que contiene todo el diseño web, opuesto, a su vez, al back-end, que es el componente de la web referido a los datos y al software para su comunicación. El front-end es de gran relevancia en tanto representa una identidad corporativa de una empresa en la Internet, algo que no es sencillo si se considera que el front-end de una página web de una compañía puede atraer o repeler clientes. Además, este elemento suele agrupar tres lenguajes diferentes de programación: el HTML, el CSS y el JavaScript, descritos a continuación (Villalba, 2015).

3.4. Lenguajes de programación HTML, JavaScript y CSS.

Como se indicó, estos tres lenguajes de programación suelen ser los más utilizados para la creación y caracterización del front-end de una página web. El

lenguaje HTML en específico, siglas correspondientes a HyperText Markup Language, es el lenguaje más usado para elaborar páginas web.

Ahora, de acuerdo con el Neowiki (neoattack.com, s.f.), el HTML no es en realidad un lenguaje que esté pensado para programar, sino que usa etiquetas para ordenar y dar forma al sinnúmero de portales de la web. Por otro lado, el lenguaje denominado CSS se usa para definir un estilo y un aspecto a documentos escritos en lenguajes como el HTML. Lo anterior significa que el lenguaje CSS dota al documento de determinados colores, tipo y tamaño de fuentes e incluso determina el espaciado entre letras y párrafos de dicho documento. Finalmente, el lenguaje JavaScript “es un lenguaje de programación que permite llevar a cabo actividades tanto simples como complejas en páginas web. Es un sistema que goza de la peculiaridad de no necesitar compilación alguna, ya que son los propios navegadores los que se encargan de leer el código para asimilarlo y llevar a cabo las acciones que este indica. A día de hoy es el estándar cuando se quieren plantear mecanismos más complicados de lo habitual en una web (neoattack.com, s.f.).

3.5. Base de datos.

Cuando se habla de una base de datos, se hace referencia a un conjunto de datos almacenados sin redundancias innecesarias en un soporte informático y accesible simultáneamente por distintos usuarios y aplicaciones (Cobo, 2007). Los datos deben ser estructurados y almacenados de forma totalmente independiente de las aplicaciones que la utilizan (Cobo, 2007).

Dicho esto, a continuación se enumeran los componentes de una base de datos de acuerdo con Cobo (2007):

- **Datos.** Es el componente fundamental de la base de datos, como se decía en la definición están relacionadas entre sí formando un conjunto con mínimas redundancias. Los datos por sí mismos no aportan conocimiento, hay que procesarlos y transformarlos.
- **Software de Sistema Gestor de Base de Datos -SGBD-.** Un sistema de gestión de base de datos es un software o conjunto de programas que permite crear y mantener una base de datos. El SGBD actúa como interfaz entre los programas de aplicación (Usuarios) y el sistema operativo. El objetivo principal del SGBD es proporcionar un entorno eficiente a la hora de almacenar y recuperar información de la base de datos. Este software facilita el proceso de definir, construir y manipular bases de datos para diversas operaciones.

Luego, se resalta que definir una base de datos consiste en especificar los tipos de los datos, las estructuras de los datos y las restricciones de los datos. Construir una BD es el proceso de almacenar los datos en algún medio de almacenamiento controlado por el SGDB una vez definida la base de datos. La manipulación de la BD se refiere a la consulta de los datos para obtener cierta información, la actualización de la BD, y la generación de informes a partir de los datos almacenados.
- **Usuarios.** Otro componente de un sistema de BD son los usuarios. Existen tres tipos diferentes de usuarios:
 - Programadores de aplicación: Escriben programas de aplicación que utilizan la base de datos.

- Usuarios “ingenuos”: Usuarios poco experimentados que utilizan las aplicaciones escritas previamente.
- Usuarios casuales: Utilizan consultas formuladas en un lenguaje de consulta de BD.
- **Administrador de la base de datos (DBA).** Es la persona o grupo de personas encargadas del control del sistema. Las funciones del DBA incluyen las siguientes:
 - Definir y modificar esquemas de la base de datos y las restricciones de los datos.
 - Crear y modificar estructuras de almacenamiento físicas y los métodos de acceso.
 - Autorizar el acceso a la BD a los usuarios.
 - Garantizar el funcionamiento correcto del sistema y prestar servicio técnico: Se ocupa principalmente de problemas de seguridad y problemas en la velocidad del sistema.
 - Realizar copias de seguridad o back -up del contenido de la base de datos.

3.6. Normalización de bases de datos.

En el análisis de datos, es de suma importancia emprender acciones que permitan mejorar el desempeño de las bases de datos evitando la redundancia en la información y ofreciendo un diseño confortable para el cliente (Coordinación de Universidad Abierta y Educación a Distancia-UNAM [CUAED], s.f.). Ante esta necesidad, la normalización de bases de datos se entiende como “la transformación de las vistas de usuario complejas

y del almacén de datos a un juego de estructuras más pequeñas y estables” (Kendall y Kendall, 2005). Por tanto, la normalización de bases de datos facilita la manipulación de grandes volúmenes de información, simplifica la extracción de datos minimizando errores lógicos y disminuye el espacio ocupado en el disco repercutiendo en la eficiencia, la eficacia y la productividad en cualquier organización (Kendall y Kendall, 2005; MySQL Hispano, 2010).

De manera similar, Espinosa (2016) indica que la normalización de bases de datos es un proceso que clasifica relaciones, objetos y formas de relación con base en las características que cada uno posee, donde sí se identifica el cumplimiento de ciertas reglas preestablecidas se aplica una categoría, o en caso contrario, se descomponen repartiendo sus atributos entre esquemas de relación más pequeños que cumplen las condiciones establecidas.

En suma, Coronel, Morris y Rob (2011) explican que la normalización de bases de datos puede ser usada en diferentes momentos, siendo; el diseño y creación de una nueva base de datos, el análisis de relaciones entre atributos después de ser creada, el análisis para el mejoramiento y optimización del diseño, y la reestructuración o modificación de una base de datos ya existente.

3.7. Otros conceptos importantes.

Por último, se mencionan algunos frameworks necesarios para la ejecución del proyecto.

Django, es un framework web de código abierto escrito en **Python** que permite construir aplicaciones web más rápidamente y con menos código. Fue desarrollado inicialmente para gestionar aplicaciones web de páginas orientadas a noticias de

WorldOnline, y más tarde se liberó bajo licencia BSD. Es de anotar que Django respeta el modelo de vista controlador. Contiene un conjunto de componentes que permite desarrollar sitios web de manera más fácil y rápida (Molina et al., 2016).

Web scraping, también conocido como Web harvesting o Web data extraction, es el proceso de rastreo y descarga de sitios web de información y la extracción de datos no estructurados o poco estructurados a un formato estructurado. Para lograrlo, se simula la exploración humana de la World Wide Web, ya sea por implementación de bajo nivel del protocolo de transferencia de hipertexto, o la incorporación de ciertos navegadores web (Tadeo et al., 2015).

4. Metodología

A continuación, se presenta la metodología adoptada para el desarrollo del proyecto en cuestión, describiendo el proceso realizado durante el mismo y las diferentes etapas y fases contempladas para dar cumplimiento a los objetivos propuestos.

4.1. Metodología de trabajo.

Se implementaron las metodologías ágiles de desarrollo Scrum teniendo presente que:

- Adoptan una estrategia de desarrollo incremental, en lugar de la planificación y ejecución completa del producto.
- Se basan en la calidad del resultado más en el conocimiento tácito de las personas en equipos auto organizados, que en la calidad de los procesos empleados.
- Evitan solapamiento de las diferentes fases del desarrollo, en lugar de realizar una tarea tras otra en un ciclo secuencial o en cascada.

Ahora, debido a que Scrum es un modelo de referencia que define un conjunto de prácticas y roles, en la Tabla 1 se muestran los roles principales definidos e implementados.

Tabla 1. Roles definidos para la implementación de la metodología Scrum.

Rol	Descripción
Cliente	Recibe el producto y puede influir en el proceso, entregando sus ideas o comentarios respecto al desarrollo. En nuestro caso, este rol lo ejerce la vicepresidencia de auditoría interna.
Product Owner o propietario	Representa la voz del cliente. Se asegura de que el equipo trabaje de forma adecuada desde la perspectiva del negocio. Es 'el jefe' responsable del proyecto. Para el proyecto, el propietario del producto es la gerencia de la sección.
Scrum Master o facilitador de proyectos	Es la figura que lidera los equipos en la gestión ágil de proyectos. Su misión es que los equipos de trabajo alcancen sus objetivos hasta llegar a la fase de “sprint final”, eliminando cualquier dificultad que puedan encontrar en el camino. Es el encargado de hablar con el Product Owner para desarrollar el producto conforme a sus especificaciones. Este papel lo asume el auditor líder de desarrollo.
Equipo	Ejecuta el desarrollo y demás elementos relacionados con él. El practicante y los

	demás auditores responsables del desarrollo conforman el equipo
--	---

Fuente: Elaboración propia.

4.2. Proceso de desarrollo.

Durante cada “sprint”, definido como un periodo entre una y cuatro semanas (a tiempo acordado por el equipo), siendo esta lo más corta posible, el equipo creó una serie de tareas potencialmente entregables (utilizable). El Product Owner, en el Sprint Planning, identificó las tareas que quería ver completadas/finalizadas y se las mostró al equipo. Entonces, el equipo conversó con el Product Owner buscando la claridad y la magnitud adecuadas, para posteriormente determinar la cantidad de ese trabajo con la que podrían comprometerse para ser completada durante el “sprint”.

La plataforma usada para el proyecto es llamada “Asana”. En esta plataforma, se pueden organizar las tareas por el equipo, añadir un tiempo a cada tarea, normalmente por el Scrum Master, y además unos puntos de esfuerzo que permitieran dar una idea de la carga laboral del desarrollador a la semana.

4.3. Etapas del proyecto.

Para efectos de este proyecto, se definieron y desarrollaron las etapas de trabajo que se presentan a continuación (Ver Tabla 2), teniendo presentes todos los requerimientos en cuanto a conocimientos y actividades a realizar para que el proyecto se culminara con éxito.

Tabla 2. Etapas definidas para el desarrollo del proyecto

Etapa	Descripción
Formación	Realización del curso ‘Diseñando páginas web con Bootstrap 4’

Diseño	Diseño de las diferentes pantallas con las que va a contar la aplicación MAR.
Estudio	Realización del estudio sobre la organización del proyecto.
Creación del proyecto	Creación del proyecto con sus respectivas carpetas y bibliotecas de ayuda.
Inicio del proyecto	Inicialización del proyecto usando GitHub.
Desarrollo e implementación	Implementación y desarrollo de las diferentes vistas diseñadas para la plataforma MAR.
Creación de las bases de datos	Creación de las bases de datos.
Desarrollo de scripts	Desarrollo de los scripts requeridos para la extracción de la información usando Python.
Análisis de resultados	Análisis de resultados obtenidos tras el diseño, el desarrollo y la implementación de las vistas de la plataforma MAR, así como de las bases de datos y los scripts para la extracción de información de los usuarios según lo planteado.

Fuente: Elaboración propia

4.4. Fases del estudio previo.

4.4.1. Diseños previos.

En este apartado, se presentan los diferentes diseños de las vistas o *mockups* propuestos para este proyecto. Los diseños fueron desarrollados por la diseñadora encargada mediante el uso del software InVision, aunque se resalta que durante el desarrollo y la producción de las vistas es posible que sufrieran leves modificaciones.

Para una mayor adaptabilidad a los dispositivos, se consideró pertinente empezar la creación de todos los diseños para las versiones móviles. Posteriormente, se desarrollaron los diseños para móviles en un formato ancho para tabletas y para ordenadores o pantallas grandes. En el caso específico de la Figura 1, se realizaron vistas similares para que al cambiar de dispositivo la página se ajustara de acuerdo con el ancho de la pantalla. Adicionalmente, se destaca que el motivo principal por el cual se planteó y se llevó a cabo dicha estrategia para la creación de los diseños de las vistas, fue la dificultad de adaptar todos los campos en una pantalla de tamaño pequeña.

Por otra parte, en este proyecto se usó el prototipo de rejillas de Bootstrap para una mayor adaptabilidad y sencillez a la hora de implementar el código.

En seguida, se presentan los diseños de las diferentes secciones que se plantearon para la plataforma MAR de Bancolombia. En la Figura 1, se muestra la vista para el home o la página de inicio, en las Figuras 2 y 3 se muestran las vistas para las secciones de Login o inicio de sesión tanto para el ingreso del usuario como para el ingreso de la contraseña, y en la Figura 4 se presenta la vista diseñada para la página correspondiente a la recuperación de contraseña.

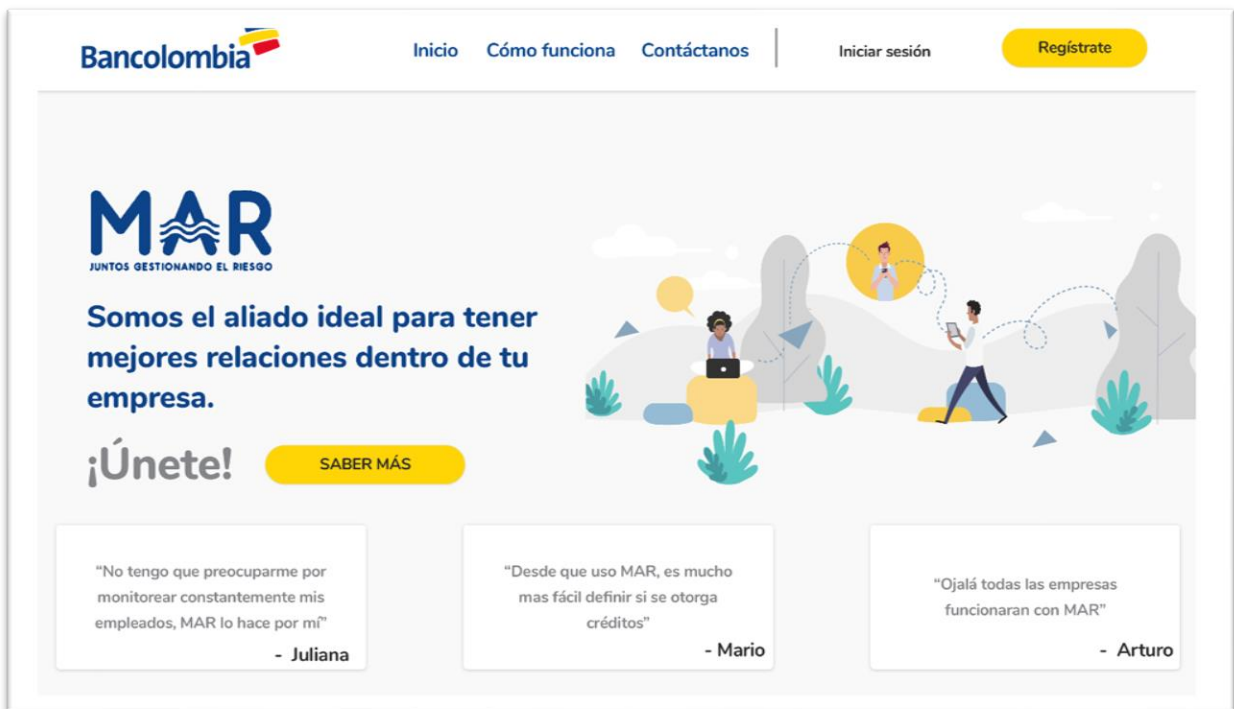


Figura 1. Vista del home o página principal de la plataforma MAR sin inicio de sesión.

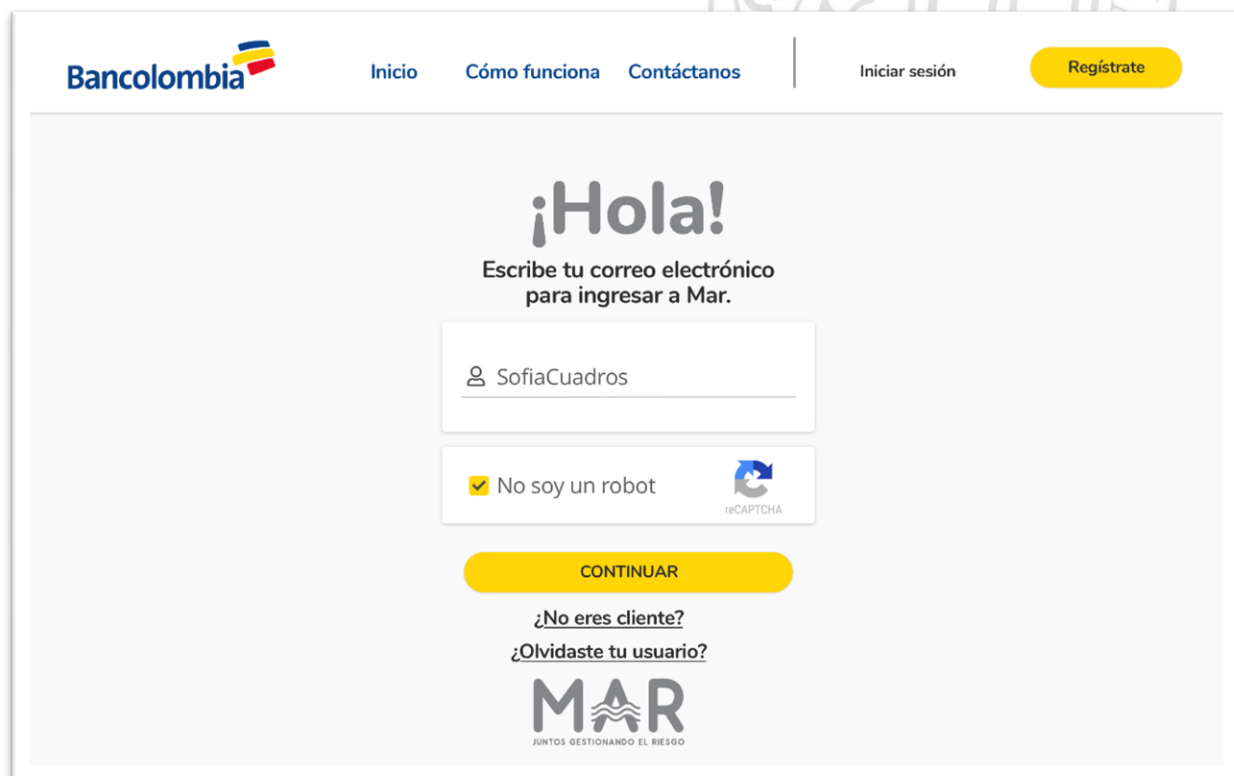


Figura 2. Vista diseñada para la página de Login o inicio de sesión, ingreso del usuario registrado en la plataforma MAR.

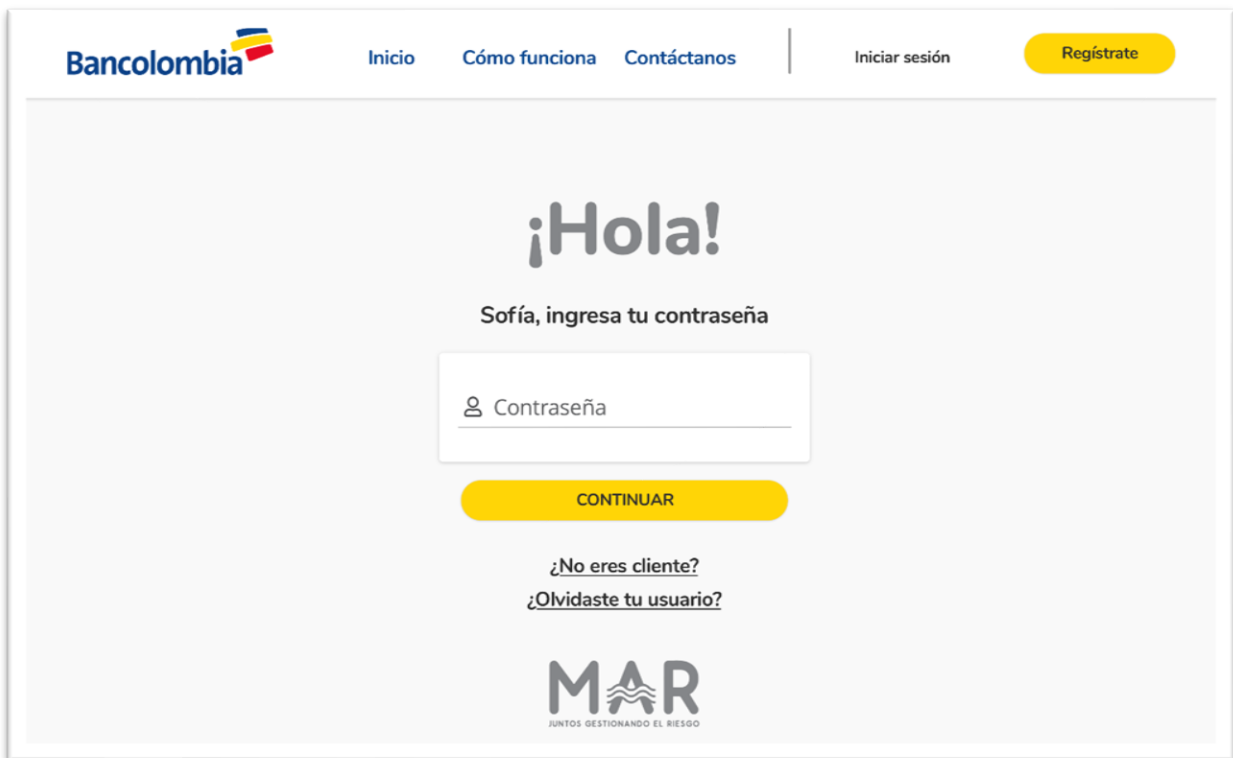


Figura 3. Vista diseñada para la página de Login o inicio de sesión, ingreso de la contraseña de la cuenta personal registrada en la plataforma MAR.

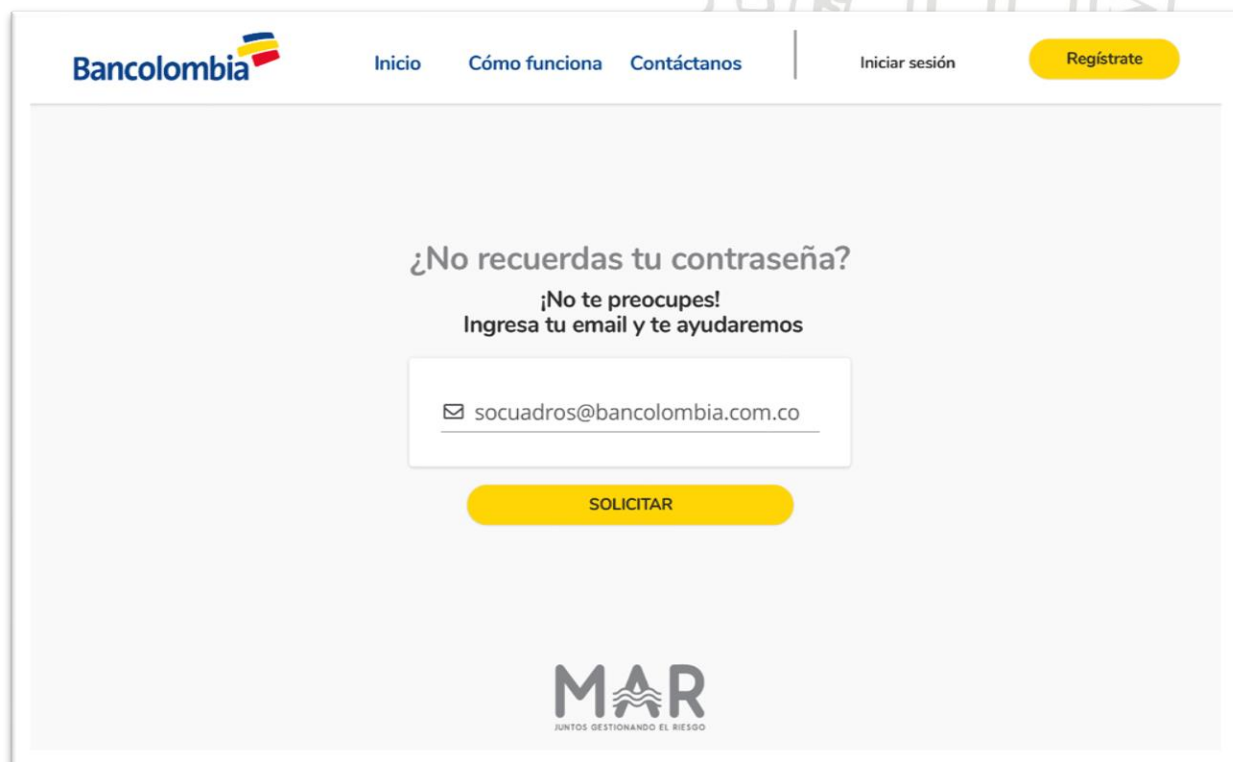


Figura 4. Vista diseñada para la página de recuperación de contraseña para usuarios registrados en la plataforma MAR.

Ahora, en las figuras 5 a 10 que se presentan a continuación, se pueden observar los diseños desarrollados para la plataforma MAR una vez el usuario ha iniciado sesión. Se muestran tanto las vistas de la página principal de inicio, como del menú desplegado, las búsquedas de perfiles de riesgo, los resultados y los historiales esencialmente, que hacen parte de la experiencia que los clientes tendrán al hacer uso de la plataforma MAR de Bancolombia.

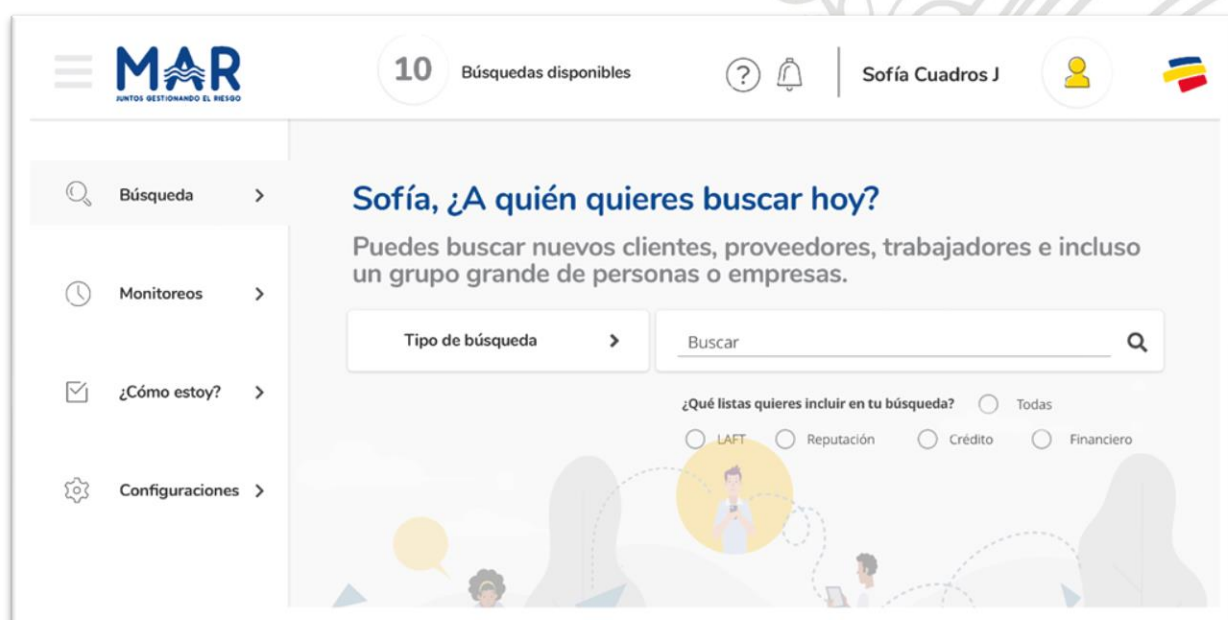


Figura 5. Vista diseñada para la página principal después del inicio de sesión en la plataforma MAR.

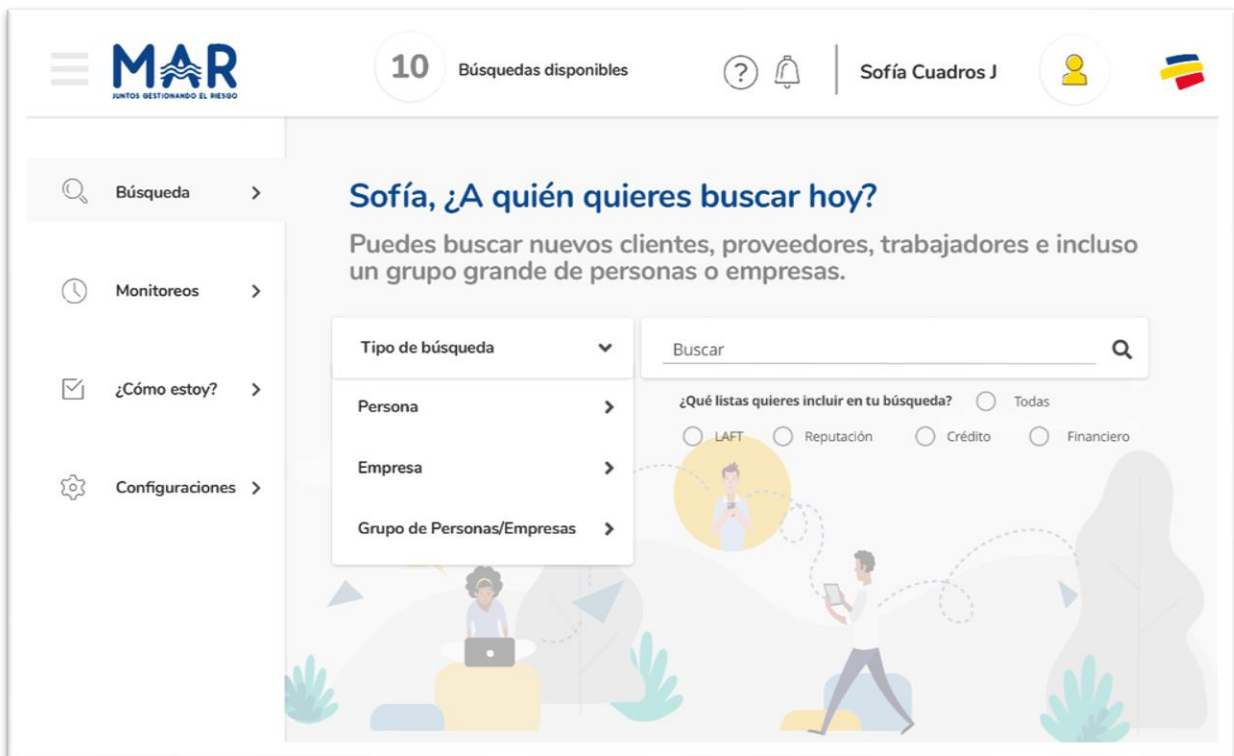


Figura 6. Vista diseñada para la página principal con el menú desplegado para definir el tipo de búsqueda después del inicio de sesión en la plataforma MAR.

The screenshot shows the MAR platform interface. At the top, there is a navigation bar with the MAR logo, a search count of 9, and the user name 'Sofía Cuadros J'. Below the navigation bar, the breadcrumb trail reads 'Búsqueda > Resultados > Detalles'. The main heading states 'Hemos encontrado los siguientes resultados:'. The primary result card for 'PLÁSTICOS S.A.S' includes general information such as NIT, address, and legal form. To the right, an 'Análisis MAR' section indicates a high-risk status. Below the main card, four detailed analysis cards are displayed: LAFT (Lavado de Activos y Financiación del Terrorismo), Reputación, Crédito, and Financiero. Each card provides a summary of risks and key metrics, with a 'CONOCER MÁS' button for further details.

Información general:

NIT: 247934
 Razón Social: PLÁSTICOS S.A.S
 Dirección: Carrera 33 No. 43 - 89, Sabaneta, Antioquia.
 Teléfono: (3) 3841048
 Dirección email: plasticos@plasticos.com.co
 Representante legal: Castro Parra, Edgar

Actividad: G435 - Comercio al por mayor de productos químicos de uso agropecuario
 Forma Jurídica: Sociedad por Acciones Simplificada
 Fecha Constitución: 12/12/1977
 Socios: No se registran
 Establecimientos: 5

Análisis MAR
Empresa de alto riesgo
 La persona/empresa está relacionada directamente con actividades de Lavado de Activo y Financiación del Terrorismo.

LAFT
 Lavado de Activos y Financiación del Terrorismo
 La persona/empresa está relacionada directamente con actividades de Lavado de Activo y Financiación del Terrorismo.

Reputación
 fecha de actualización: Dic 2019
 Se recomienda realizar un mayor conocimiento del cliente, presenta antecedentes significativos que pueden impactar tu empresa.

Crédito
 fecha de actualización: Dic 2019
 Es recomendable hacer una gestión adecuada de su cartera comercial al establecer alguna relación comercial con esta empresa.

Financiero
 fecha de actualización: Dic 2019
 Es recomendable solicitar algún tipo de garantías al establecer alguna relación comercial con esta empresa.

Bloqueos:	1
Alertas:	0
Prensa negativa:	1

Sanciones:	1
Certificaciones:	1
Prensa Positivas:	1
Prensa Negativas:	1

Días en mora: <small>Últimos 24 meses</small>	20
Créditos concluidos: <small>Últimos 9 años</small>	3
Cupo asignado:	\$3.000.000

Activo:	\$24.000.000
Pasivo:	\$30.000.000
Patrimonio:	\$28.600.000
Ventas:	\$22.000.000
Utilidades:	\$4.000.000

Figura 7. Vista diseñada para la página de resultados de búsqueda dentro de la plataforma MAR.



Búsquedas

Porcentaje de concidencia:



80%

Para disminuir o aumentar la cantidad de resultados mostrados, filtra los resultados por porcentaje de coincidencia. Entre más alto sea el porcentaje, más precisa es tu búsqueda

Listas favoritas:

Para disminuir o aumentar la cantidad de listas analizadas filtra los resultados según tus listas de preferencia

LAFI

- Todas las fuentes
- OFAC
- ONU
- DEA
- Black List Ucrania
- Sanciones Financieras Europa
- Interpol
- Sanciones Autónomas Canadá
- CSL
- BM - Anticorrupción
- Sanciones del IADB
- Lista Duque/ FTO
- Cadáveres CNI
- Sanciones ejecutorias
- Most Wanted ATF
- Unauthorised Firms

Crédito

- Data Crédito

Financiero

- Estados Financieros

Reputación


Sanciones

- Todas las fuentes
- Boletín de Deudores Morosos del Estado
- Sanciones Superfinanciera
- Sanciones Autorregulados del Mercado de Valores de Colombia
- Contadores y sociedades prestadoras de servicios contables con sanciones
- CPAE Consejo Administración Empresas
- CPIP Consejo profesional Ingeniería Petroleo
- CPIB Consejo profesional de Biología
- Boletín de Sanciones SIC

Certificados

- Todas las fuentes
- Empresas B

<input type="radio"/>	La República	<input type="radio"/>	ABC
<input type="radio"/>	Caracol	<input type="radio"/>	Neue Zürcher Zeitung
<input type="radio"/>	The New York Times	<input type="radio"/>	Süddeutsche Zeitung
<input type="radio"/>	The Times	<input type="radio"/>	Primera Hora
<input type="radio"/>	New York Post	<input type="radio"/>	El Vocero
<input type="radio"/>	USA Today	<input type="radio"/>	El Mercurio
<input type="radio"/>	Folha Do Brasil	<input type="radio"/>	La República Uruguay
<input type="radio"/>	Correio Braziliense	<input type="radio"/>	La Jornada
<input type="radio"/>	Venezuela Analysis	<input type="radio"/>	El País
<input type="radio"/>	Los tiempos	<input type="radio"/>	Tal Cual
<input type="radio"/>	Prensa Libre	<input type="radio"/>	Hoy
<input type="radio"/>	El siglo	<input type="radio"/>	El Diario de Cuba
<input type="radio"/>	Diario Libre	<input type="radio"/>	El Comercio
<input type="radio"/>	Diario de Occidente	<input type="radio"/>	La Tribuna
<input type="radio"/>	The Globe and Mail	<input type="radio"/>	Siglo XXI
<input type="radio"/>	La Repubblica Italia	<input type="radio"/>	Neue Zürcher Zeitung
<input type="radio"/>	Frankfurter Allgemeine Zeitung	<input type="radio"/>	Aftonbladet
<input type="radio"/>	Le Monde	<input type="radio"/>	The Press and Journal
<input type="radio"/>	Le Figaro	<input type="radio"/>	Blind
<input type="radio"/>	Liberation	<input type="radio"/>	Página 12

 **Notificaciones**

Correos:
 Registra los correos de las personas a las que quieres que automáticamente se comparta información de monitoreos.

Acciones:
 Selecciona las notificaciones que quieres que sean enviadas

Ejecución del monitoreo
 Nuevas listas de búsqueda
 Alertas encontradas

GUARDAR

Figura 8. Vista diseñada para la página de configuraciones dentro de la plataforma MAR.

MAR 10 Búsquedas disponibles | Sofía Cuadros J

Sofía, ¿Conoces las regulaciones necesarias para tu empresa en Sostenibilidad y LAFT (Lavado de activos y Financiación del terrorismo)?

LAFT

Porcentaje de cumplimiento: **40%**

Identifica cómo está tu sistema de regulación LAFT dentro de la empresa.

CONOCER MÁS

Sostenibilidad

Porcentaje de cumplimiento: **20%**

Identifica cómo está tu sistema de regulación Sostenible dentro de la empresa.

CONOCER MÁS

LAFT

Autodiagnóstico: 40% | Plan de trabajo: 20% | Mis archivos: 2

Políticas	Identificación de riesgos	Herramientas tecnológicas	Documentos y registros	Procedimientos	Capacitación	
				Sí	No	Adjuntar
01	Compromiso de cero tolerancia ¿La Junta Directiva ha declarado su compromiso de cero tolerancia al lavado de activos y la financiación del terrorismo?			<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="text" value=""/>
02	Definición de deberes ¿Cuenta con políticas que definen los deberes de los órganos administrativos, oficial de cumplimiento y de todos los miembros de la organización?			<input type="checkbox"/>	<input checked="" type="checkbox"/>	
03	Políticas de operaciones ¿Cuenta con políticas para abstenerse de operaciones con personas que no estén plenamente identificadas?			<input type="checkbox"/>	<input checked="" type="checkbox"/>	
04	Obligaciones de funcionarios ¿Cuenta con políticas respecto a la obligación de los funcionarios de colaborar con autoridades competentes?			<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="text" value=""/>
05	SARLAFT ¿En los procedimientos definidos se consignan las sanciones por incumplimiento a las normas relacionadas con el SARLAFT, así como los procesos para su imposición?			<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="text" value=""/>

01 - 05 de 25 preguntas < >

Figura 9. Vista diseñada para la página de consulta *¿Cómo estoy?* dentro de la plataforma MAR.

Figura 10. Vista diseñada para la página de *Historial de búsqueda y Monitoreos* dentro de la plataforma MAR.

4.4.2. Arquitectura del proyecto.

En esta sección de la fase previa, se diseñó la arquitectura del proyecto, que consiste en la creación del directorio de carpetas y en una primera estructura de la base de datos. En la siguiente imagen, Figura 11, se puede observar el directorio general del proyecto. A continuación, se detallan las carpetas más importantes del proyecto:

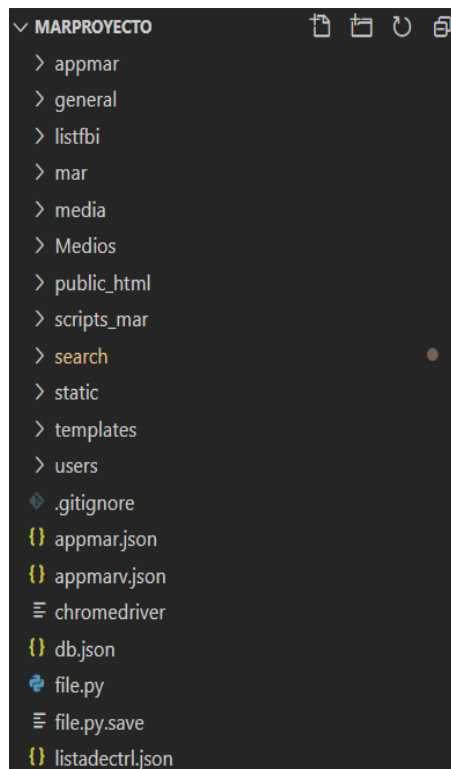


Figura 11. Estructura del directorio general del proyecto.

Con respecto a la estructura del directorio general del proyecto que se ilustra en la Figura 11, se describen en la Tabla 3 los contenidos de algunas de las carpetas establecidas debido a su relevancia.

Tabla 3. Contenido de algunas carpetas establecidas dentro de la estructura del directorio general del proyecto.

Carpeta	Descripción
mar	Contiene un pequeño script para calcular el tiempo en que se hace logout por inactividad.
media	Contiene los recursos usados en las vistas de imágenes, tipos de fuente, etc.
Medios	Contiene los scripts que hace el Web scraping de los medios de comunicación.

scripts_mar	Contiene los scripts para la extracción de información de las listas de control.
search	Contiene las URL de cada una de las vistas de la plataforma MAR que se generan de manera automática con Django.
templates	Es la carpeta más importante, pues contiene los códigos HTML de cada una de las vistas diseñadas y a usar en la plataforma MAR.

4.5. Fase de producción.

4.5.1. Front-end.

Como se presentó en el marco teórico, cuando se habla del front-end se hace referencia a los elementos de la página con los que pueden interactuar los usuarios. Dicho esto, desde las vistas en la pantalla de login y de registro, hasta las barras de carga y los mensajes de alerta deben aparecer cada vez que se realiza alguna petición. Por este motivo, en este apartado se profundiza en los elementos con los que el usuario puede interactuar en cada pantalla, ya que hay que tener en cuenta que cada formulario y cada campo requiere una validación para evitar que la base de datos guarde valores no deseados o en su defecto, evitar el surgimiento de problemas asociados a la seguridad de la plataforma en general.

A manera de ejemplo, en la Figura 12 que se muestra a continuación, se encuentra el fragmento del código HTML (front-end) construido que corresponde a la vista de la página de inicio de sesión en la plataforma MAR de la Figura 13.

```

</div>
<div class="col-12 col-xl-12 col-lg-12 col-md-12 col-sm-12 colorbackgr">
  <div class="container-fluid">
    <div class="d-flex justify-content-center">
      <div class="bloque text-center">
        <div id="loadinglogin" class="hidden">
          
        </div>
        <div class="contenidologin">
          <div class="contenttext">
            <h1 class="">¡Hola!</h1>
            <p id="logmessage">Ingresa tu correo electrónico.</p>
          </div>
          <form id="loginmar" class="formufin" method="POST">
            {% csrf_token %}
            <div class="form-group">
              <div class="loginblock ">
                <input id="user" required class="form-control logininput" type="email" name="user" placeholder="Ingresa tu correo electrónico">
              </div>
            </div>
            <span class="text-danger" id="catcha_error" ></span>
            <div class="g-recaptcha customcatpcha" data-callback="enableBtn" data-expired-callback="unableBtn" data-size="normal" data-sitekey="{{site_key}}">
            </div>
            <button type="submit" id="btn-log" class="btn btn-login" disabled="disabled">CONTINUAR</button>
          </form>
        </div>
        <div class="pt-4 redirectlog">
          <p> <a href=" ../users/registro">¿No eres cliente?</a></p>
          <p><a href=" ../users/reset_password">¿Olvidaste tu usuario o contraseña?</a></p>
          
        </div>
      </div>
    </div>
  </div>
</div>
</div>

```

Figura 12. Código HTML



Figura 13. Visualización del código Front-end para el inicio de sesión en la plataforma MAR.

4.5.1.1. Acciones y visualizaciones en las vistas.

Como se ha venido discutiendo, las vistas diseñadas originalmente pueden tener pequeñas diferencias con respecto a las desarrolladas, razón por la cual se resalta que las vistas que no estén marcadas con la numeración anterior al *mockup* es porque tuvieron pequeños cambios durante la fase de producción.

Tabla 4. Descripción de las vistas y los elementos que contiene

Página a visualizar	Descripción
Home o página principal sin inicio de sesión (Figura 1)	Se puede acceder a las siguientes vistas: <ul style="list-style-type: none"> - Saber más. - Inicio (Figura 1). - ¿Cómo funciona?

	<ul style="list-style-type: none"> - Contáctanos. - Iniciar sesión, ingreso de usuario (Figura 2) - Regístrate.
<p>Login o inicio de sesión, ingreso del usuario registrado (Figura 2)</p>	<p>Se puede ingresar el usuario e ir a vista para ingresar la contraseña (Figura 3). Asimismo, se puede acceder a las vistas</p> <ul style="list-style-type: none"> - ¿Olvidaste tu usuario? (Figura 4) Si no recuerda su usuario o contraseña. - ¿No eres cliente? Si no está registrado el usuario. - Inicio (Figura 1), incluyendo ¿Cómo funciona? y Contáctanos. <p>Finalmente, también incluye el CATCHAP que debe resolver el usuario para iniciar sesión.</p>
<p>Login o inicio de sesión, ingreso de la contraseña (Figura 3)</p>	<p>En esta vista, es posible ingresar la contraseña para ir a la vista principal de la plataforma después de inicial sesión (Figura 5). Luego, se puede acceder desde ésta a otras vistas como:</p> <ul style="list-style-type: none"> - ¿Olvidaste tu usuario? (Figura 4), si no recuerda su contraseña - ¿No eres cliente?, Si el cliente no está registrado en la plataforma. - Retornar al Home (Figura 1), incluyendo las vistas de ¿Cómo funciona? y Contáctanos.
<p>Recuperar contraseña (Figura 4)</p>	<p>Es posible diligenciar el campo de email para así reestablecer su usuario o contraseña. En suma, se puede acceder a las vistas de:</p> <ul style="list-style-type: none"> - Inicio (Figura 1). - ¿Cómo funciona? - Contáctanos. - Iniciar sesión (Figura 2) - Regístrate

<p>Página principal (inicio de sesión)</p>	<p>Esta vista consta de una barra en la parte superior, tomada de la documentación de Bootstrap, en donde se puede visualizar el logo de MAR, un contador que nos indica el número de búsquedas restantes en el plan de pago, el nombre y la foto de perfil del usuario en sesión, que al darle clic nos llevará a la vista de perfil de usuario y un botón que, desplegará una barra lateral para que el usuario se desplace por diferentes vistas como:</p> <ul style="list-style-type: none"> - Normatividad - Monitoreo (Figura 10). - Búsqueda. - Configuraciones. <p>En la parte central se encuentra un menú, que permite al usuario escoger el <i>tipo de búsqueda</i> que desea, por persona o empresa. A la derecha de este menú, una barra de búsqueda en la que se ingresa el nombre de la persona o empresa a consultar, según sea el caso.</p> <p>Finalmente, contiene 4 checkbox para caracterizar la búsqueda y tomando un enfoque únicamente en los tipos de listas seleccionados.</p>
<p>Resultados de búsqueda (Figura 7)</p>	<p>Como se mencionó en la vista anterior, las barras laterales de perfil y navegación están presentes. Ahora, en la vista de los resultados se encuentra la información concerniente al nombre de la empresa o persona que está consultando el usuario. Se desplegará la información general:</p> <ul style="list-style-type: none"> - Nit o cédula de la empresa o persona buscada. - Nombre o razón social de la empresa. - Dirección. - Teléfono. - Email.

- Representante legal.
- Actividad comercial en la cual se desempeña la empresa o persona.
- Forma jurídica.
- Fecha de constitución.
- Socios.
- Establecimientos.

Asimismo, se presentará la **calificación**, para lo cual se definieron 5 *cards* en donde está contenida la calificación final del cliente y un botón que lleva al cliente a unos datos más detallados de la búsqueda. Cuatro de estas *cards* son la calificación individual de cada uno de las categorías definidas (LAFT, Reputación, Crédito y Financiero), y una última *card* se ha asignado para la calificación global ponderada a partir de los 4 puntajes asociados a cada lista de consulta.

A continuación, se detallará la información contenida en cada una de las *cards*:

LAFT

- Bloqueos o número de listas en las que se encuentra contenido.
- Número de alertas.
- Prensa negativa.

Reputación

- Sanciones.
- Certificaciones.
- Prensa positiva.
- Prensa negativa.

Crédito

- Días de mora.
- Créditos concluidos.

	<ul style="list-style-type: none"> - Cupo asignado. <p>Financiero</p> <ul style="list-style-type: none"> - Activo. - Pasivo. - Patrimonio. - Ventas. - Utilidades. <p>Análisis MAR</p> <ul style="list-style-type: none"> - Calificación final. - Descripción de la calificación. - Botón que permite descargar el informe
Configuraciones (Figura 8)	<p>En la parte superior de esta vista, se puede encontrar un botón con el que se puede reiniciar la configuración, dado el caso en que el usuario haya realizado algún cambio. De manera similar, se dispone de una barra con un porcentaje de coincidencia, es decir, al hacer variaciones de 0 a 100 el sistema hará algo diferente con la búsqueda: si la barra está al 100% únicamente encontrará resultados de personas o empresas en donde el nombre ingresado en el inicio sea exactamente igual. Por otro lado, si este porcentaje empieza a bajar, el sistema no será tan estricto y ampliará el rango de búsqueda con algunos caracteres comunes.</p> <p>Más abajo, se encuentra un conjunto de listas clasificadas dependiendo de su origen - LAFT, Crédito, Financiero y Reputación - con un pequeño checkbox que al ser marcado indicará qué se desea buscar en esta lista. Por defecto, todas están marcadas y el usuario puede desmarcar las listas que no desee consultar.</p>

	<p>Por último, se encuentra un edit text en el cual el usuario puede colocar su email para recibir la información de configuración a su correo electrónico.</p>
Normatividad	<p>En la vista de normatividad, el cliente puede consultar cómo está su empresa respecto a los 4 ítems principales de consulta de riesgo ya mencionados, y un botón en la parte inferior que lo redirecciona para ver resultados más detallados.</p>
Monitoreo	<p>En esta vista, se presentan los monitoreos que el cliente tenga. Al hacer una búsqueda y encontrar un resultado, el cliente puede seleccionar la opción de monitoreo, que hace un seguimiento en el tiempo al resultado de búsqueda, permitiéndole informarse sobre algún cambio en la calificación asignada.</p> <p>También, se dispone de un conjunto de botones en la parte inferior: Activar, Desactivar, Actualizar, Descargar y Borrar, donde cada uno tiene una función correspondiente a su nombre.</p> <p>Finalmente, esta vista contiene un botón llamado <i>nuevo monitoreo</i> en donde el usuario podrá iniciar nuevos seguimientos a las búsquedas de su interés.</p>

Adicionalmente, se deja explícito que se integró el código con el proyecto principal, para lo cual se hizo una sincronización por medio *Git* para que los demás desarrolladores tuvieran disponible el nuevo código y así poder realizar el ensamble con el desarrollo back-end; se realiza un *merge* con el apoyo de un desarrollador encargado de la plataforma.

4.5.2. Base de datos.

En este apartado se detallan la estructura y el diseño de la base de datos, al igual que los campos que contiene cada una de las tablas establecidas. Para el proyecto desarrollado, se tuvieron en cuenta 5 tablas principales que están encargadas de almacenar la información de las personas o empresas previamente buscadas, las cuales son alimentadas a medida que se van realizando las consultas en la plataforma MAR con el fin de evitar la realización del mismo proceso desde la web originaria cada vez que se realice una consulta similar y de evitar redundancias dentro de la misma base de datos. Lo anterior, constituye una estrategia para realizar consultas más rápidas a la BD.

En la Figura 14, se muestra el esquema con la estructura establecida para la base de datos, en el cual se pueden discriminar las 5 tablas mencionadas para optimizar el flujo de información.

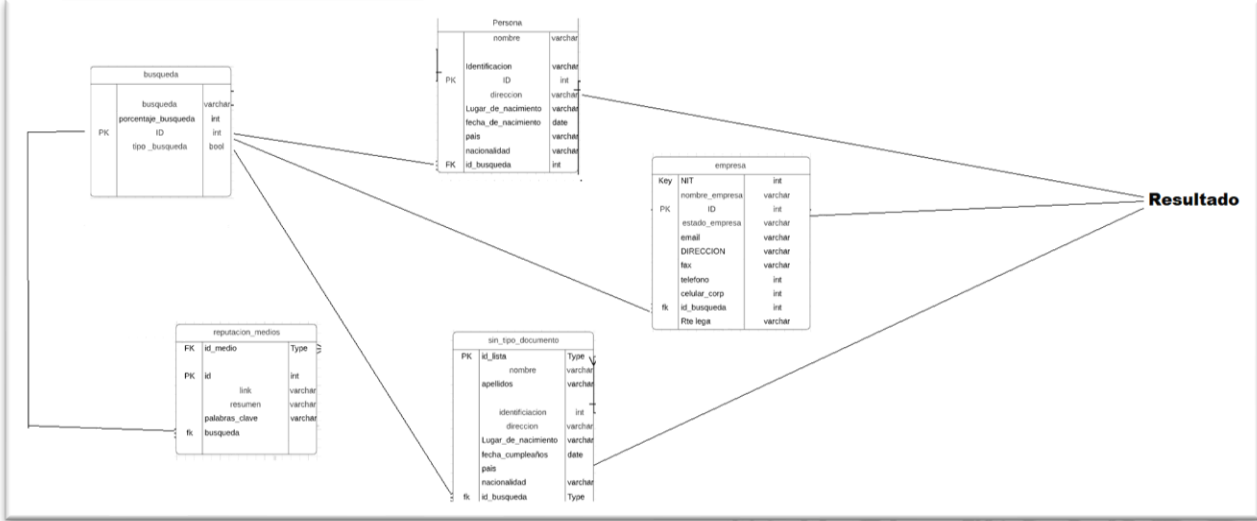


Figura 14. Diseño base de datos. **Fuente:** Elaboración propia

En ese sentido, en la Tabla 5 se presenta la descripción detallada de cada una de las 5 tablas planteadas que conforman la base de datos.

Tabla 5. Descripción de las tablas que conforman la base de datos

Tabla	Descripción
-------	-------------

Tabla de Búsqueda	<p>Esta tabla es la encargada de comenzar la búsqueda en la base de datos, dependiendo de la búsqueda que realice el cliente desde la vista de inicio. Sus campos son:</p> <ul style="list-style-type: none"> - Búsqueda: Tipo varchar, máximo 250 caracteres. - Porcentaje_búsqueda: Tipo tinyint, tamaño de 0-100. - ID: Tipo tinyint, primary key, tamaño de 0-65535. - Tipo_búsqueda: Tipo bool, valores de 0-3.
Tabla de Personas	<p>En esta tabla se encuentra la información personal de las personas que son encontradas en las listas LAFT, crédito, financiero y reputación mediante la técnica de Web scraping ejecutada por el desarrollador, de la cual hablaremos más adelante.</p> <ul style="list-style-type: none"> - id: Tipo tinyint, primary key, tamaño de 0-65535. - id_búsqueda: Tipo tinyint, foreign key, tamaño de 0-65535. - Identificación: Tipo varchar, tamaño máximo 50 caracteres. - Nombre: Tipo varchar, tamaño máximo 50 caracteres. - Dirección: Tipo varchar, tamaño máximo 30 caracteres. - lugar_de_nacimiento: Tipo varchar, tamaño máximo 30 caracteres. - fecha_de_nacimiento: Tipo date. - País: Tipo varchar, tamaño máximo de 0 30 caracteres. - Nacionalidad: Tipo varchar, tamaño máximo de 0 30 caracteres.
Tabla de empresas	<p>En esta tabla se encuentra la información personal de las empresas que son encontradas en las listas LAFT, crédito, financiero y reputación mediante la técnica de Web scraping ejecutada por el desarrollador, de la cual hablaremos más adelante.</p>

	<ul style="list-style-type: none"> - id: Tipo tinyint, primary key, tamaño de 0-65535. - id_búsqueda: Tipo tinyint, foreign key, tamaño de 0-65535. - Nit: Tipo varchar, tamaño máximo 50 caracteres. - nombre_empresa: Tipo varchar, tamaño máximo 50 caracteres. - estado_empresa: Tipo varchar, tamaño máximo 30 caracteres. - Email: Tipo varchar, tamaño máximo 30 caracteres. - Dirección: Tipo varchar, tamaño máximo 30 caracteres. - Fax: Tipo varchar, tamaño máximo 30 caracteres. - Teléfono: Tipo varchar, tamaño máximo 30 caracteres. - celular_corp: Tipo varchar, tamaño máximo 30 caracteres. - representante_legal: Tipo varchar, tamaño máximo 30 caracteres. - Actividad: Tipo varchar, tamaño máximo 30 caracteres. - forma_juridica: Tipo varchar, tamaño máximo 30 caracteres. - fecha_constitucion: Tipo date. - Socios: Tipo varchar, tamaño máximo 30 caracteres. - Establecimiento: Tipo varchar, tamaño máximo 30 caracteres.
Tabla sin_tipo_doc.	En esta tabla se encuentra la información de las personas o empresas que no tienen especificado si su número de identificación es un NIT o una cédula.

	<ul style="list-style-type: none"> - Id: Tipo tinyint, primary key, tamaño de 0-65535. - Nombre: Tipo varchar, tamaño máximo 50 caracteres. - Identificación: Tipo varchar, tamaño máximo 50 caracteres. - Dirección: Tipo varchar, tamaño máximo 30 caracteres. - Lugar_de_nacimiento: Tipo varchar, tamaño máximo 30 caracteres. - fecha_de_nacimiento: Tipo varchar, tamaño máximo 30 caracteres. - País: Tipo varchar, tamaño máximo 30 caracteres. - Nacimiento: Tipo date. - id_búsqueda: Tipo tinyint, foreign key, tamaño de 0-65535.
Reputación medios	<p>En esta tabla se descarga la información obtenida de los medios de comunicación físicos o digitales establecidos por el Product Owner.</p> <ul style="list-style-type: none"> - ID_medio: Tipo tinyint, foreign key, tamaño de 0-65535. - Id: Tipo tinyint, primary key, tamaño de 0-65535. - Link: Tipo varchar, tamaño máximo 300 caracteres. - Resumen: Tipo varchar, tamaño máximo 2000 caracteres. - Palabra_clave: Tipo varchar, tamaño máximo 50 caracteres. - Búsqueda: Tipo tinyint, foreign key, tamaño de 0-65535.

4.5.2.1. *Funcionamiento del Web scraping*

En la Figura 15, se muestra la página web asociada a la lista de control de la Policía Nacional de Colombia, la cual se toma como ejemplo ilustrativo para dar evidencias sobre el funcionamiento del Web scraping, ya que esta hace parte del conjunto de listas de interés para realizar consultas en la plataforma MAR. En esta lista de control se observa con claridad un listado de los directores y de los comandantes de la institución indicando el rango, el nombre y la unidad a la que pertenece. La verificación de los elementos de la página web consultada se realizó con ayuda de la herramienta de desarrollador de Google y la extensión XPath Helper.

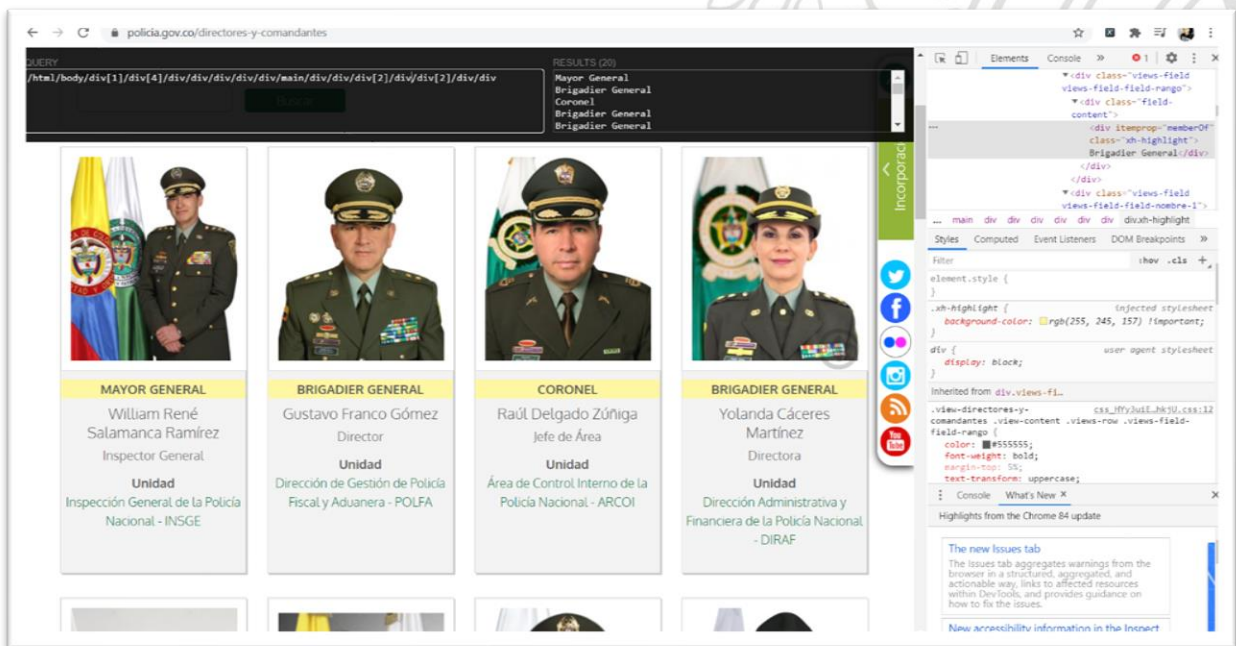


Figura 15. Vista de la página de la Policía Nacional de Colombia

En seguida, se muestra la Figura 16 en la cual se presenta la identificación por medio de la herramienta XPath Helper de cada elemento listado donde se puede evidenciar el rango de todos los funcionarios de la policía que aparecen en esta página; esto se realiza por medio del XPath y la estructura HTML, que permiten obtener el texto asociado a cada uno de los elementos, en este caso es el rango.

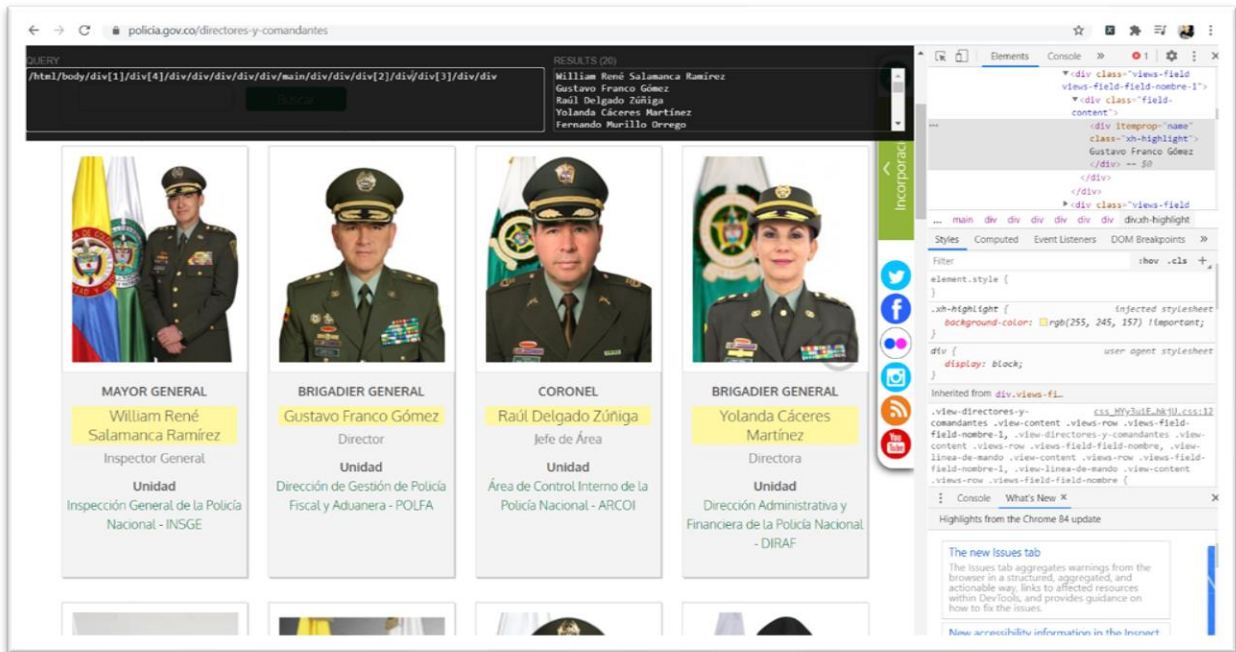


Figura 16. Lista nombre con elementos Xpath y la estructura HTML

Después, en la Figura 17 se presenta la identificación por nombre de cada elemento listado donde se puede ver dicho texto asociado a todos los funcionarios de la policía que aparecen en esta página.

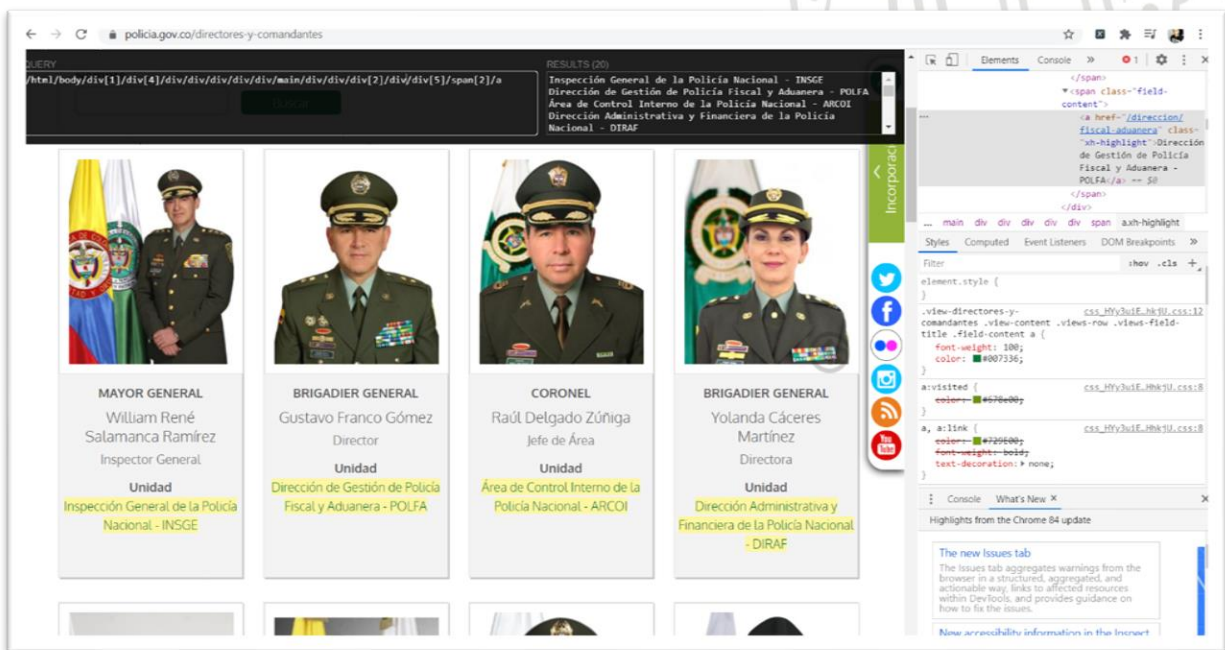


Figura 17. Lista unidad con elementos Xpath en la estructura HTML

Luego, con el mismo análisis realizado y presentado en las Figuras 16 y 17, se procedió a obtener la unidad a la que pertenece cada funcionario de la policía con la estructura XPath obteniendo una lista con el nombre de todas las unidades correspondientes a cada miembro de la policía.

En la Figura 18, se presentan las rutas XPath en Python por medio de las cuales se referencian las listas creadas con la información extraída de la página web de interés, en este caso, en la página de la Policía Nacional (nombre, rango y unidad de los funcionarios).

```
Lista_Rangos=driver.find_elements_by_xpath('/html/body/div[1]/div[4]/div/div/div/div/div/main/div/div/div[2]/div/div[2]/div/div')
Lista_Nombre=driver.find_elements_by_xpath('/html/body/div[1]/div[4]/div/div/div/div/div/main/div/div/div[2]/div/div[3]/div/div')
Lista_Unidad =driver.find_elements_by_xpath('/html/body/div[1]/div[4]/div/div/div/div/div/main/div/div/div[2]/div/div[5]/span[2]/a')
```

Figura 18. Búsqueda de elementos Xpath en Python

Con base a las librerías investigadas, se utilizó el método ‘find_elements_by_xpath’ (Muthukadan, 2018), donde se le da como parámetro el XPath identificado en las figuras anteriores. De esta forma, el script puede encontrar los elementos de la página web y puede guardarlos en un vector. La Figura 19 muestra un fragmento de la matriz que guarda los resultados obtenidos en Python después de realizar el Web scraping en la página de la Policía Nacional. En esta misma figura, se observa que el tamaño de la matriz obtenida es 101 (filas) x 3 (columnas), donde cada columna corresponde a los vectores (o listas) de elementos de interés: cargo, nombre y unidad a la que pertenece el funcionario.

	Rango	Nombre_Funcionario	Unidad a la que pertenece
0	MAYOR GENERAL	William René Salamanca Ramírez	Inspección General de la Policía Nacional - INSGE
1	BRIGADIER GENERAL	Gustavo Franco Gómez	Dirección de Gestión de Policía Fiscal y Aduan...
2	CORONEL	Raúl Delgado Zúñiga	Área de Control Interno de la Policía Nacional...
3	BRIGADIER GENERAL	Yolanda Cáceres Martínez	Dirección Administrativa y Financiera de la Po...
4	BRIGADIER GENERAL	Fernando Murillo Orrego	Dirección Antisecuestro y Antiextorsión de la ...
..
96	CORONEL	Edilberto García Guauta	Departamento de Policía Vichada - DEVIC
97	CORONEL	Juan Carlos Castellanos Álvarez	Oficina de Comunicaciones Estratégicas - COEST
98	TENIENTE CORONEL	Yasid Alberto Montaña Granados	Escuela Antidrogas "Mayor Wilson Quintero Mart...
99	CORONEL	Alba Patricia Lancheros Silva	Unidad Policial para la Edificación de la Paz
100	TENIENTE CORONEL	Carlos Antonio Ardila Rocha	Área de Relaciones y Cooperación Internacional...

[101 rows x 3 columns]

Figura 19. Resultados Web scraping visualizado en Python

Por último, se especifica que este proceso se repitió con cada una de las listas de interés dadas por el banco para alimentar la plataforma MAR. Aproximadamente 53 listas fueron obtenidas.

4.5.2.2. Funcionamiento de la base de datos.

Después de haberse extraído la información desde la página web y guardarse en una tabla, el script de Python consume desde la vista de inicio producción el porcentaje de coincidencia y el texto de búsqueda. Luego, internamente por medio del método *SequenceMatcher* de la librería *difflib*, se calcula el valor de coincidencia del texto de búsqueda con los valores internos de los datos extraídos, y si este valor es mayor al configurado por el cliente, se procede a guardar estos datos en un DataFrame de Pandas llamado “resultados”, cuya información será visible en la vista resultados de la plataforma MAR.

Seguidamente, se verifica que la tabla “resultados” no esté vacía y se procede a llamar la función *InsertarEnBusqueda*. En esta función se insertan los datos correspondientes a porcentaje de coincidencia, búsqueda, tipo de búsqueda y un id único generado incrementalmente por Django a la tabla de búsqueda y retorna el máximo id

que haya en la tabla con el fin de verificar cuál registro es el último y cuál se está consultando actualmente.

Por último, se procede a insertar a la tabla personas, empresas o sin_tipo_doc todos los registros de la tabla resultados, más el id de búsqueda de la siguiente manera (Ver Figura 20). La finalidad de la tabla “personas” y “Empresas” es que cuando el cliente o un cliente diferente deseen realizar la búsqueda con parámetros **similares** (Nombre de búsqueda y porcentaje de similitud) , el **SGDB** no deba realizar el Web scraping nuevamente , crear el DataFrame resultados y almacenarlos en listas individuales, sino que directamente consultará la tabla personas o empresa, haciendo el cálculo de similitud que encuentre en cada uno de ellos y mostrándolo directamente en la vista resultados, acelerando en aproximadamente 15 segundos la consulta. De manera análoga, se procede con la tabla de medios.

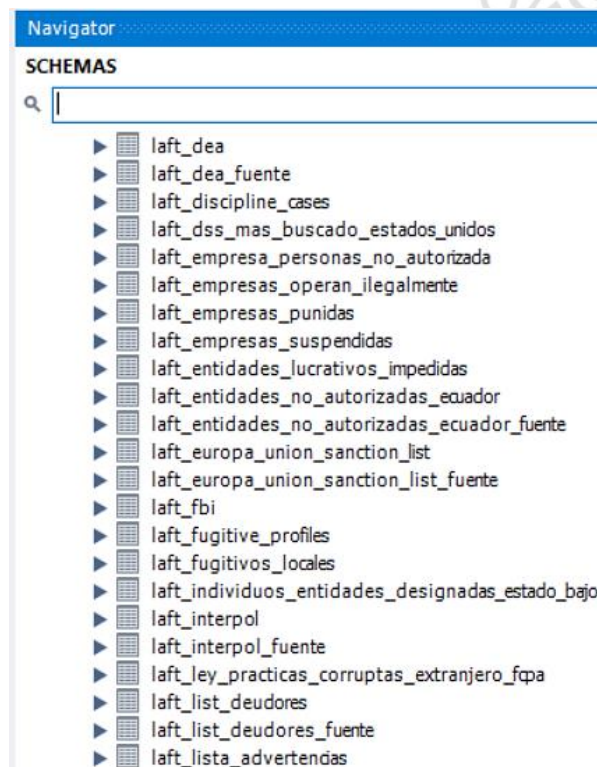


Figura 20. Tablas de las listas en el Workbench

La Figura 21 es un ejemplo de los resultados obtenidos en Tabla Persona, después de poner en funcionamiento la base de datos ante una búsqueda.

nombre	identificacion	id	direccion	lugar_nacimiento	fecha_de_nacimiento	pais	nacionalidad	id_búsqueda
Carlora perea	254693	1	cra 57	medellin	25/07/1995	colombia	colombiano	1
Carla perez	2478561	2	cra 48	lima	22/06/1997	argentina	peruano	1
Carlos pelaez	458789	3	cra 47	bucaramanga	22/01/1996	colombia	colombiano	1

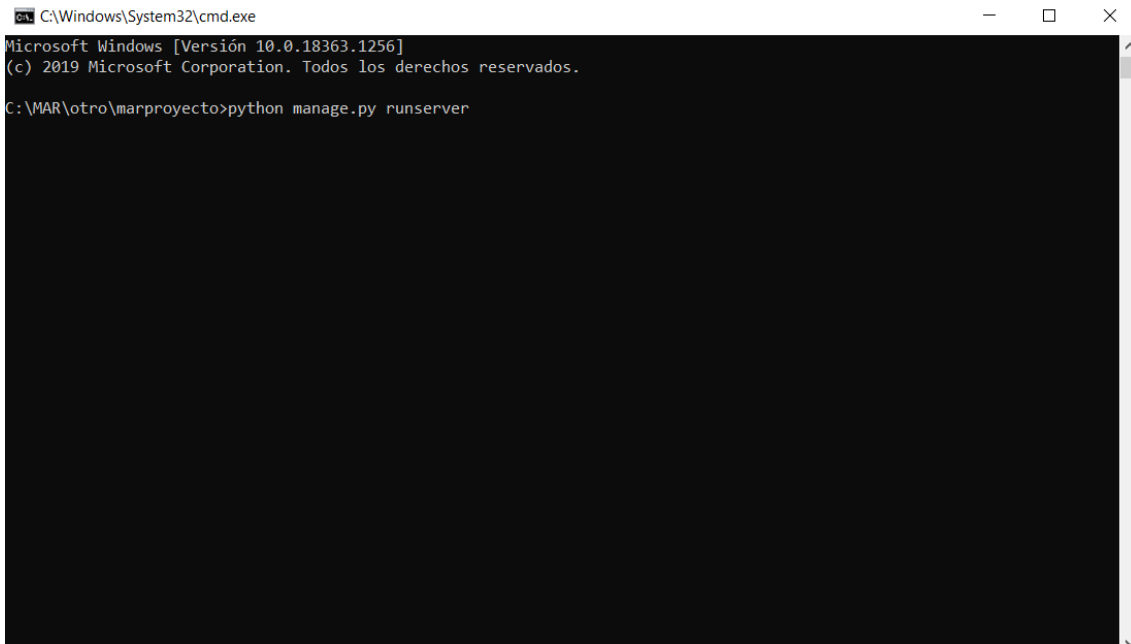
Figura 21. Ejemplo de resultados obtenidos en tabla persona

5. Análisis de resultados

El resultado obtenido es una plataforma web en forma de producto mínimo viable -PMV. El funcionamiento del software fue testeado semana a semana en cada uno de los cierres de sprint por un auditor experto en el tema que participó en el desarrollo y validó su funcionalidad parcial, pues al ser un PMV aún queda camino por recorrer.

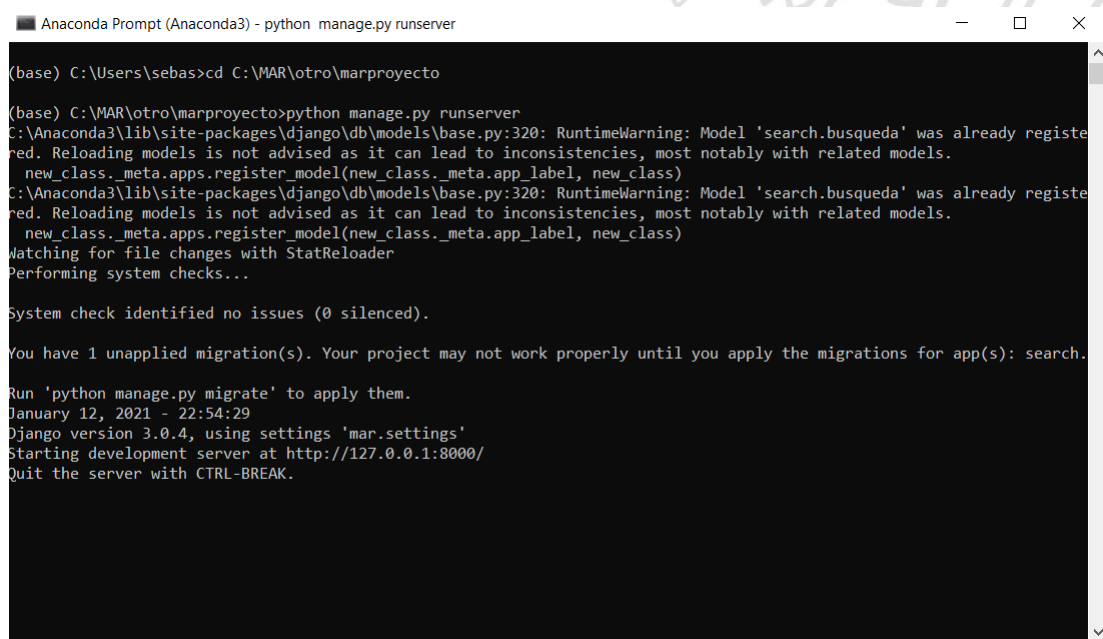
En las Figura 22 se muestra el resultado final del producto, presentando cada una de las vistas desarrolladas para la plataforma MAR de Bancolombia y las bases de datos con registros con consultas realizadas a modo de prueba.

Para empezar, se mostrará el proceso de ingreso a la plataforma de manera local aun sin la respectiva traducción de nombre de dominio.



```
C:\Windows\System32\cmd.exe
Microsoft Windows [Versión 10.0.18363.1256]
(c) 2019 Microsoft Corporation. Todos los derechos reservados.
C:\MAR\otro\marproyecto>python manage.py runserver
```

Figura 22. Comando para inicializar el servidor.



```
Anaconda Prompt (Anaconda3) - python manage.py runserver
(base) C:\Users\sebas>cd C:\MAR\otro\marproyecto
(base) C:\MAR\otro\marproyecto>python manage.py runserver
C:\Anaconda3\lib\site-packages\django\db\models\base.py:320: RuntimeWarning: Model 'search.busqueda' was already registered. Reloading models is not advised as it can lead to inconsistencies, most notably with related models.
  new_class._meta.apps.register_model(new_class._meta.app_label, new_class)
C:\Anaconda3\lib\site-packages\django\db\models\base.py:320: RuntimeWarning: Model 'search.busqueda' was already registered. Reloading models is not advised as it can lead to inconsistencies, most notably with related models.
  new_class._meta.apps.register_model(new_class._meta.app_label, new_class)
Watching for file changes with StatReloader
Performing system checks...

System check identified no issues (0 silenced).

You have 1 unapplied migration(s). Your project may not work properly until you apply the migrations for app(s): search.
Run 'python manage.py migrate' to apply them.
January 12, 2021 - 22:54:29
Django version 3.0.4, using settings 'mar.settings'
Starting development server at http://127.0.0.1:8000/
Quit the server with CTRL-BREAK.
```

Figura 23. Dirección IP a la plataforma MAR

Como se puede ver en la Figura 23, se generó una dirección IP número 127.0.0.1 con puerto 8000.



Figura 24. Home sin inicio de sesión resultado de producción

Se puede evidenciar una diferencia considerable entre la Figura 24 producto y la Figura 1 de diseño realizado por la diseñadora web en la zona de comentarios, más específicamente en los tamaños de las fuentes, esto producto a cambios de diseño realizados en el camino.

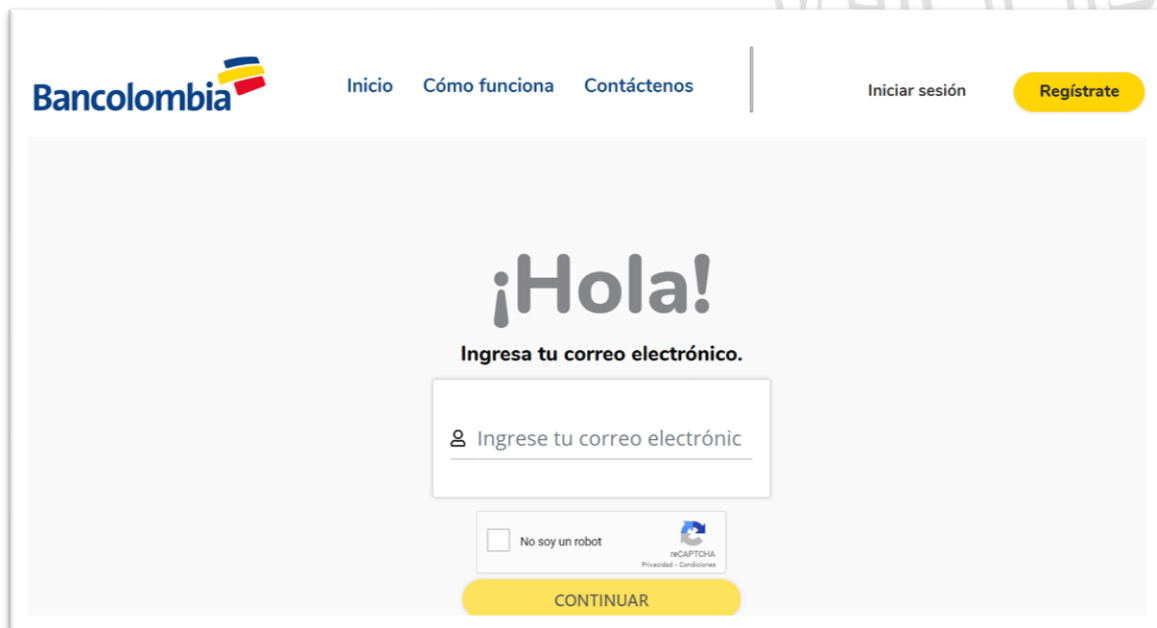
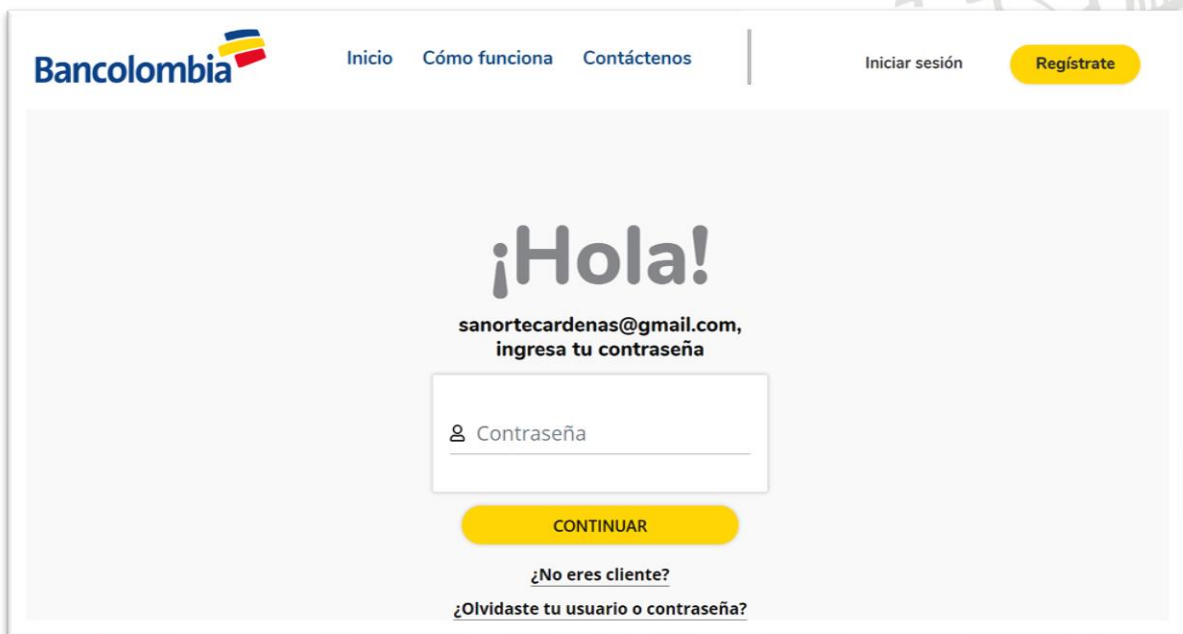


Figura 25. Inicio de sesión resultado de producción

La Figura 25, en comparación con la Figura 2 que correspondía al diseño construido de la vista se observó una diferencia en lo que se refiere a los tamaños de las fuentes usadas. Esto mismo ocurre en lo que respecta a las vistas de inicio de sesión para el ingreso de contraseña y la vista de la página para la recuperación de contraseña para usuarios previamente registrados; las Figuras 26 y 27 muestran las vistas producidas.



Bancolombia Inicio Cómo funciona Contáctenos Iniciar sesión [Regístrate](#)

¡Hola!

sanortecardenas@gmail.com,
ingresa tu contraseña

[CONTINUAR](#)

[¿No eres cliente?](#)
[¿Olvidaste tu usuario o contraseña?](#)

Figura 26. Inicio sesión para el ingreso de contraseña en la plataforma MAR resultado de la producción

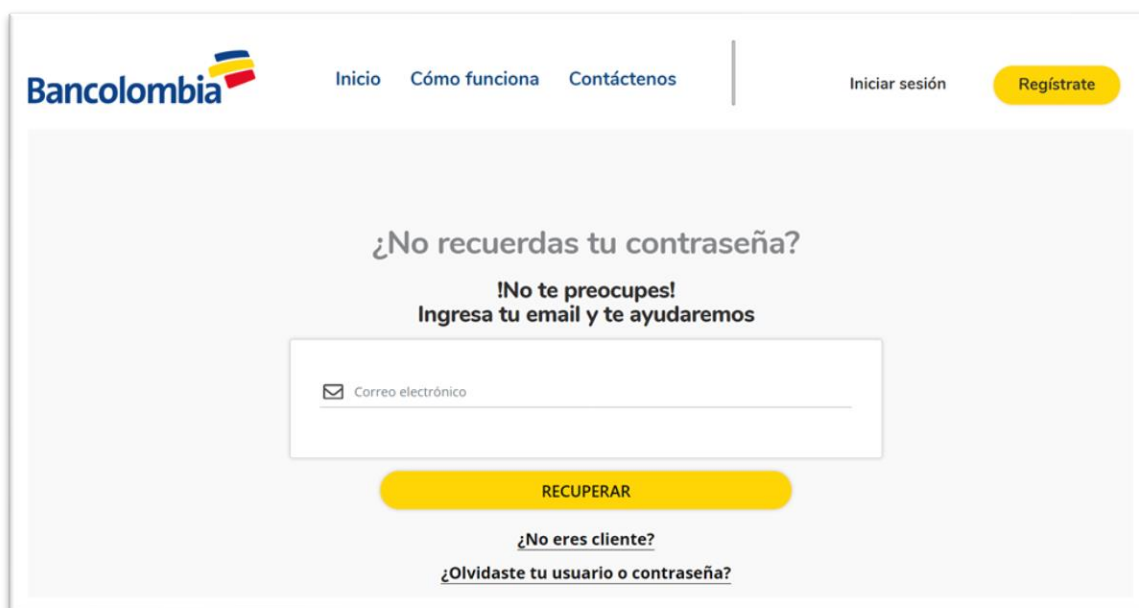


Figura 27. Vista recuperar contraseña para usuarios registrados resultado de la producción

Ahora, en la Figura 28 que se presenta a continuación se muestra la vista obtenida de la página principal de la plataforma MAR con inicio sesión después de ser producida. Con respecto al diseño planteado en la Figura 5, se encuentra que la vista producida difiere del checkbox original en el cual se pintaban los tipos de listas disponibles para búsqueda.



Figura 28. Vista de la página principal de la plataforma con inicio de sesión, resultado de la producción

Como parte de análisis de resultados, se considera relevante dejar explícito que, con respecto al diseño original de la vista de la sección de resultados encontrados tras realizar una búsqueda en la plataforma MAR de Bancolombia (Figura 7), la vista producida no tiene ningún cambio; esta vista obtenida se presenta en la Figura 29.

MAR 9 Búsquedas disponibles Santiago Ortega

Búsqueda > Resultados > Detalles

Hemos encontrado los siguientes resultados:

Información general:
 Nit: 247934
 Razón Social: PLÁSTICOS S.A.S
 Dirección: Carrera 33 No. 43 - 89, Sabaneta, Antioquia.
 Teléfono: (3) 3841048
 Dirección email: plasticos@plasticos.com.co
 Representante legal: Castro Parra, Edgar

Actividad: G435 - Comercio al por mayor de productos químicos de uso agropecuario
 Forma Jurídica: Sociedad por Acciones Simplificada
 Fecha Constitución: 12/12/1977
 Socios: No se registran
 Establecimientos: 5

Análisis MAR
 Empresa de alto riesgo
 La persona/empresa está relacionada directamente con actividades de Lavado de Activo y Financiación del Terrorismo.
 MONITOREAR
 DESCARGAR INFORME

LAFT
 Lavado de Activos y Financiación del Terrorismo
 La persona/empresa está relacionada directamente con actividades de Lavado de Activo y Financiación del Terrorismo.
 Bloqueos: 1
 Alertas: 0
 Prensa negativa: 1
 CONOCER MÁS

Reputación
 Fecha de actualización: Dic 2019
 Se recomienda realizar un mayor conocimiento del cliente, presenta antecedentes significativos que pueden impactar tu empresa.
 Sanciones: 1
 Certificaciones: 1
 Prensa Positivas: 1
 Prensa Negativas: 1
 CONOCER MÁS

Crédito
 Fecha de actualización: Dic 2019
 Es recomendable hacer una gestión adecuada de su cartera comercial al establecer alguna relación comercial con esta empresa.
 Días en mora: 20 (Últimos 24 meses)
 Créditos concluidos: 3 (Últimos 10 años)
 Cupo asignado: \$3.000.000
 CONOCER MÁS

Financiero
 Fecha de actualización: Dic 2019
 Es recomendable solicitar algún tipo de garantías al establecer alguna relación comercial con esta empresa.
 Activo: \$24.000.000
 Pasivo: \$30.000.000
 Patrimonio: \$28.600.000
 Ventas: \$22.000.000
 Utilidades: \$4.000.000
 CONOCER MÁS

Figura 29. Vista de la página de Resultados en la plataforma MAR tras realizar una búsqueda, resultado de la producción

En lo que corresponde a la vista de la página de configuraciones dentro de la plataforma MAR, se identificaron algunos cambios sustanciales en cuanto al tamaño de la fuente y de los checkbox incluidos. En la Figura 30 se presenta la vista obtenida tras su producción.

Configuraciones

Búsquedas

Porcentaje de coincidencia: 100%

Reajustar configuración

Para disminuir o aumentar la cantidad de resultados mostrados, filtra los resultados por porcentaje de coincidencia. Entre más alto sea el porcentaje, más precisas es tu búsqueda.

Listas favoritas:

Para disminuir o aumentar la cantidad de listas analizadas filtra los resultados según tu lista de preferencia.

LAF

<input type="radio"/> Todas las fuentes	<input checked="" type="checkbox"/> OFAC
<input checked="" type="checkbox"/> ONU	<input checked="" type="checkbox"/> DEA
<input checked="" type="checkbox"/> Black List Ucrania	<input checked="" type="checkbox"/> Sanciones Financieras Europa
<input type="radio"/> Interpol	<input checked="" type="checkbox"/> Sanciones Autónomas Canadá
<input type="radio"/> CSL	<input checked="" type="checkbox"/> BM Anticorrupción
<input type="radio"/> Sanciones del IADB	<input checked="" type="checkbox"/> Lista Duque FTD
<input checked="" type="checkbox"/> Cadáveres CNI	<input checked="" type="checkbox"/> Sanciones ejecutorias
<input checked="" type="checkbox"/> Most Wanted ATF	<input checked="" type="checkbox"/> Unauthorised Firms

Crédito Data Crédito

Financiero Estados Financieros

Reputación

Sanciones Todas las fuentes

- Boletín de Deudores Morosos del Estado
- Sanciones Superfinanciera
- Sanciones Autorregulados del Mercado de Valores de Colombia
- Contadores y sociedades prestadoras de servicios contables con sanciones
- CPAE Consejo Administración Empresas
- CPIP Consejo profesional Ingeniería Petrolero
- CPIP Consejo profesional de Biología
- Boletín de Sanciones SIC

Certificados Todas las fuentes

- Empresas B
- Certificados SENA

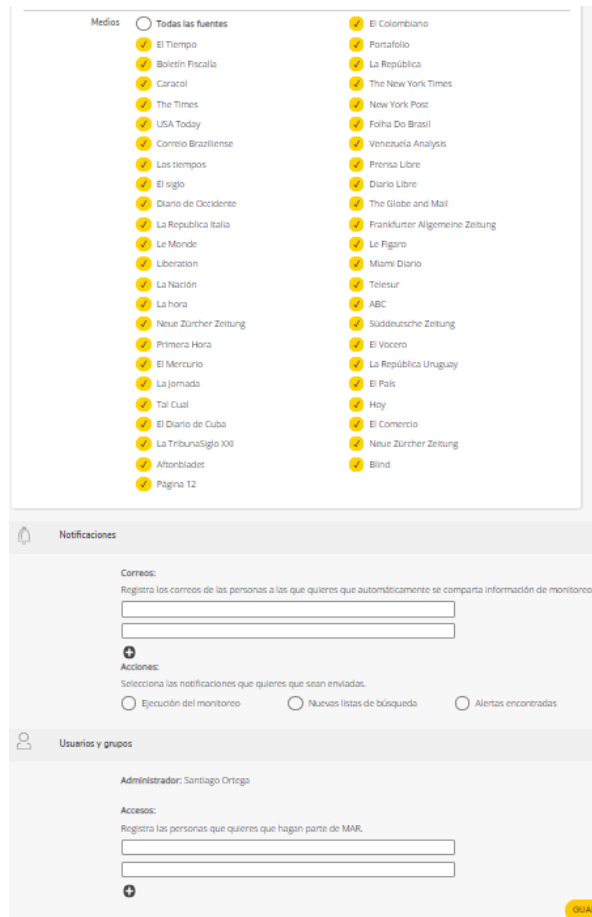


Figura 30. Vista de la página de configuración dentro de la plataforma MAR, resultado de la producción

En ese mismo orden, la vista de la página de Normatividad dentro de la plataforma MAR, comparando el diseño original (Figura 9) con la vista producida (Figura 31), presenta una diferencia evidente que consiste en que para ver los resultados detallados y dar clic para solicitar esta acción, se abre una ventana nueva (Ver Figura 32) en lugar de mostrarlos en la parte inferior como se tenía previsto.



Figura 31. Vista de la página de Normatividad dentro de la plataforma, resultado de la producción

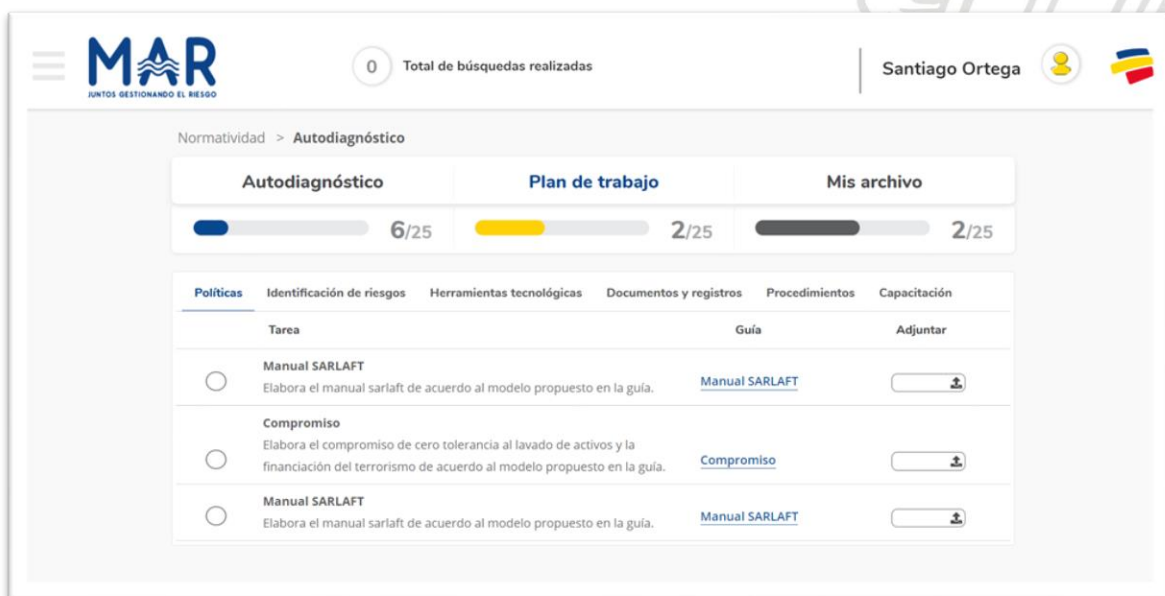


Figura 32. Vista de la nueva página de Normatividad detallada dentro de la plataforma, resultado de la producción

En última instancia, en la Figura 33 se presenta la vista producida de la página correspondiente a la sección de Historial de Búsqueda y Monitoreos. Con referencia al diseño original de esta vista planteada y mostrada en la Figura 10, se encontró que es la vista con más cambios, pues básicamente fue rediseñada cambiando su estructura y su funcionalidad.

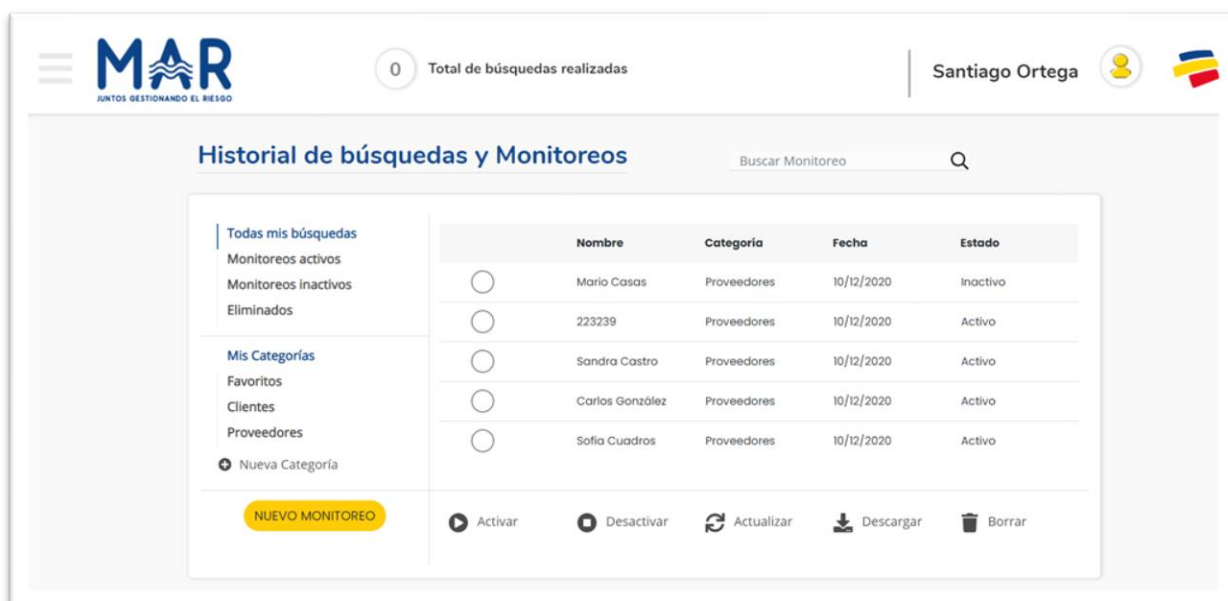


Figura 33. Vista rediseñada de la página de Historial de búsquedas y Monitoreos dentro de la plataforma MAR, resultado de la producción

La Figura 34 es un registro de los test de prueba empleados en la Tabla personas.

id	ID_BUSQUEDA	IDENTIFICACION	NOMBRE	DIRECCION	LUGAR_DE_NACIMIENTO	FECHA_DE_CUMPLEANOS
1	3	1128457081	ANDRES FELIPE MARIN HENAO	NULL	NULL	NULL
2	4	94516879	NULL	NULL	NULL	NULL
3	5	94516879	ALEXANDER VARON SANDOVAL	NULL	NULL	NULL
4	6	1152713534	FELIPE LONDOÑO MENDOZA	NULL	NULL	NULL
5	7	1007210793	JONATHAN ARISMENDY VERGARA	NULL	NULL	NULL
6	9	30338327	ADRIANA FRANCO MURCIA	NULL	NULL	NULL
7	10	8462416	HERLEY DARIO URIBE RESTREPO	NULL	NULL	NULL
8	11	8472416	MIGUEL ANGEL LONDOÑO FERNANDEZ	NULL	NULL	NULL
9	12	9103314932	Komunistyczna Partia Filipin	NULL	NULL	NULL
10	13	16377667	WILLIAM SALINAS RAMIREZ	NULL	NULL	NULL
11	14	13542606	ARNOLDO PEREZ TOLOZA	NULL	NULL	NULL
12	19	1022094598	LUIS ALFREDO BARRERA LONDOÑO	NULL	NULL	NULL
13	25	1022094598	LUIS ALFREDO BARRERA LONDOÑO	NULL	NULL	NULL
14	26	1022094598	LUIS ALFREDO BARRERA LONDOÑO	NULL	NULL	NULL
15	29	1007421506	KELLY VANNESSA	NULL	NULL	NULL
NULL	NULL	NULL	NULL	NULL	NULL	NULL

Figura 34. Registro de los testeos empleados en la tabla personas

6. Conclusiones

Las dinámicas actuales tanto a nivel social como a nivel económico suponen nuevos desafíos y exigencias para las organizaciones en todo lo referido al manejo, la gestión y el control de grandes flujos de información. Bancolombia, una entidad

bancaria con gran trayectoria y reconocimiento, ha tenido la capacidad de reconocer sus falencias y crear nuevas estrategias que les permita solventarlas en pro de la competitividad, la productividad y la eficiencia. Es por ello, que desde la gerencia de Investigación, Desarrollo e Innovación se propuso la creación y el desarrollo de una plataforma de Mercado de Autogestión de Riesgo (MAR) por medio de la cual pueden ofrecerles a sus clientes una herramienta de consulta novedosa para la prevención de riesgos.

Dicho lo anterior, en el presente informe se presentaron los resultados obtenidos posteriores a la ejecución del proyecto de soporte, acompañamiento e intervención orientado al desarrollo front –end de la plataforma MAR en el marco de la normalización de bases de datos e implementando el uso de la técnica de Web scraping. En primer lugar, se presentaron las metodologías implementadas para la extracción de información desde diferentes páginas web para alimentar la base de datos que respalda la plataforma MAR y luego, se generaron los códigos Front-end para darle lugar a las diferentes vistas que componen esta misma plataforma al ser manipulada por el usuario.

Al respecto, se destaca que aunque la mayoría de las vistas producidas solo tuvieron cambios leves con respecto a las vistas diseñadas previamente, la sección de Historial de Búsqueda y Monitoreo fue rediseñada en términos de estructura y funcionalidad pensando en las necesidades de los usuarios.

Por último, se concluye que esta primera versión construida de la plataforma MAR funciona correctamente, constituyendo una herramienta de gran utilidad tanto para Bancolombia como entidad, como para sus clientes teniendo en cuenta que se permite la agilización de procesos, evita la repetición sistemática de procesos, facilita

las consultas, reduce tiempos de acción y alimenta el portafolio de servicios ofrecidos. No obstante, se recomienda continuar con el análisis de la plataforma y la base de datos con el objetivo de realizar mejoras y optimizar al máximo los resultados en cuanto a funcionamiento y operatividad de esta herramienta.

7. Referencias

Cobo, A. (2007). Diseño y programación de bases de datos. Editorial Visión Libros: España.

Coordinación de Universidad Abierta y Educación a Distancia-UNAM [CUAED] (s.f.). Normalización de bases de datos. Unidad de apoyo para el aprendizaje. Disponible en: https://programas.cuaed.unam.mx/repositorio/moodle/pluginfile.php/872/mod_resource/content/1/contenido/index.html

Coronel, C., Morris, S. y Rob, P. (2011). Bases de datos. Diseño, implementación y administración. México: Cengage Learning.

Delgado, J. M. (2014) Desarrollo de un prototipo virtualizado de computación de alto desempeño (HPC) para la analítica de datos [Trabajo de Grado] Disponible en: <https://alejandria.poligran.edu.co/bitstream/handle/10823/1201/Trabajo%20de%20Grado%20-%20VFinal.pdf?isAllowed=y&sequence=1>

Espinosa, A. (2016). Programa Educativo de Informática Administrativa. Proceso de normalización de base de datos en una modelo relacional. Universidad Autónoma del Estado de México. Disponible en: <http://ri.uaemex.mx/bitstream/handle/20.500.11799/64072/secme16823.pdf?sequence=1>

- Fernández, H. (s.f.). ¿Qué es una RegTech y por qué los bancos acabarán utilizando esta tecnología? *Economía TIC*. Consultado el 5 de enero de 2021. Recuperado de: <https://economiatic.com/que-es-regtech/>
- Kendall, K. E. y Kendall, J. E. (2005). *Análisis y diseño de sistemas*. México: Prentice Hall.
- Masadelante.com (2019). ¿Qué es una base de datos?- Definición de base de datos. *Masadelante.com*. Consultado el 5 de enero de 2021. Recuperado de: <https://www.masadelante.com/faqs/base-de-datos>
- Molina, J. R., Loja, N. M., Zea, M. P. y Loaiza, E. L. (2016). Evaluación de los Frameworks en el desarrollo de aplicaciones web con Python. *Revista Latinoamericana de Software*, 4(4), 201-207. Disponible en: <http://revistas.unla.edu.ar/software/article/view/1149>
- Muthukadan, B. (2018). Selenium with Python. Disponible en: <https://selenium-python.readthedocs.io/>
- MySQL Hispano (2010). Normalización de bases de datos. Disponible en: <https://bustamanteg.files.wordpress.com/2010/07/normalizacion.pdf>
- Neoattack.com (s.f.). Neowiki Concepto de HTML. Neoattack.com. Consultado el 6 de enero de 2021. Recuperado de: <https://neoattack.com/neowiki/html/amp/>
- Tadeo, A., Gómez, E., Berdejo, C., Montero, J., Calderón, A. y Ibarra, R. (2015). Metodologías para análisis político utilizando Web Scraping. *Research in Computing Science*, 95, 113-121. Disponible en: https://www.rcs.cic.ipn.mx/2015_95/Metodologias%20para%20 analisis%20politico%20utilizando%20Web%20Scraping.pdf

Villalba, D. (2015). *Desarrollo de una plataforma backend y frontend para la gestión de contenidos* [Tesis de pregrado]. Universidad Carlos III de Madrid, España.

Disponible en: https://e-archivo.uc3m.es/bitstream/handle/10016/23761/TFG_Daniel_Villalba_Trujillo_2015.pdf?sequence=1&isAllowed=y

Weber, R. H. (2017). Regtech as a new legal challenge, *Journal of Financial Transformation*, 46, pp. 10-17

