

ÁREA DE TECNOLOGÍAS EMERGENTES INDUSTRIALES E

INFORMÁTICAS

PROGRAMA ACADÉMICO DE TECNOLOGÍAS DE LA INFORMACIÓN

ÁREA DESARROLLO DE SOFTWARE MULTIPLATAFORMA

METAPHORCE

PARA LA EMPRESA



FORTE INNOVATION CONSULTING

PRESENTA

JONATHAN URIEL RODRÍGUEZ HERNÁNDEZ

PARA OBTENER EL TÍTULO DE TÉCNICO SUPERIOR UNIVERSITARIO EN TECNOLOGÍAS DE LA INFORMACIÓN ÁREA DESARROLLO DE SOFTWARE MULTIPLATAFORMA

ASESORA ACADÉMICA: M.C. MARÍA DOLORES JUÁREZ RAMÍREZ

ASESOR ORGANIZACIONAL: DR. ALÁN LÓPEZ MARTÍNEZ

GENERACIÓN: 2021-2023

LEÓN, GUANAJUATO. AGOSTO 2023

RESUMEN

El proyecto realizado para la empresa Forte Innovation Consulting consistió en automatizar la búsqueda de talento para contratación en Metaphorce, con el objetivo de optimizar el proceso y ahorrar tiempo a los reclutadores.

Para ello, se realizó una adquisición de requerimientos, se diseñó la arquitectura del sistema utilizando diagramas UML y se desarrolló el software utilizando el lenguaje de programación Python.

Además, se realizó una investigación para recolectar datos en LinkedIn. Como resultado, se obtuvieron una lista de requerimientos, un diseño de arquitectura modular y la exitosa integración del software en la plataforma de Metaphorce. Este proyecto cumplió con ofrecer una solución eficiente para optimizar el proceso de contratación y mejorar la eficacia de la empresa.

AGRADECIMIENTOS

Este informe no sería posible sin la ayuda a Forte Innovation Consulting por brindar la oportunidad de realizar el proyecto en su organización. Además, se aprecia las oportunidades brindadas, fundamentales para la realización del proyecto.

Agradecer al Dr. Alan López Martínez por fungir como asesor organizacional. Siendo su ayuda de gran importancia para el desarrollo del informe final de estadía.

De igual manera agradece a María Dolores Juárez por su contribución como asesora académica. Su gran apoyo fue crucial para la elaboración y revisión del informe final de estadías.

Además de agradecer a mi compañera Lizbeth Garnica Negrete por su asistencia en la redacción del informe. Su ayuda fue importante para dar mayor claridad a algunas ideas.

**ÍNDICE**

[INTRODUCCIÓN 1](#_Toc142497198)

[1. ANTECEDENTES 3](#_Toc142497199)

[2. PROBLEMÁTICA 5](#_Toc142497200)

[3. JUSTIFICACIÓN 6](#_Toc142497201)

[4. OBJETIVOS 8](#_Toc142497202)

[4.1. OBJETIVO GENERAL 8](#_Toc142497203)

[4.2. OBJETIVOS METODOLÓGICOS 8](#_Toc142497204)

[4.3. ALCANCE 8](#_Toc142497205)

[5. METODOLOGÍA 9](#_Toc142497206)

[5.1. ANÁLISIS DE LOS REQUERIMIENTOS FUNCIONALES DEL RPA 10](#_Toc142497207)

[5.2. DISEÑO DE LA ARQUITECTURA DEL RPA PARA UNA CLARA IDENTIFICACIÓN DE REQUERIMIENTOS MEDIANTE DIAGRAMAS UML 14](#_Toc142497208)

[5.3. DESARROLLO DEL MÓDULO DE RPA PARA SU UTILIZACIÓN MEDIANTE PYTHON 18](#_Toc142497209)

[6. RESULTADOS 28](#_Toc142497210)

[6.1. ANÁLISIS DE LOS REQUERIMIENTOS FUNCIONALES DEL RPA 29](#_Toc142497211)

[6.2. DISEÑO DE LA ARQUITECTURA DEL RPA PARA UNA CLARA IDENTIFICACIÓN DE REQUERIMIENTOS MEDIANTE DIAGRAMAS UML 31](#_Toc142497212)

[6.3. DESARROLLO DEL MÓDULO DE RPA PARA SU UTILIZACIÓN MEDIANTE PYTHON 36](#_Toc142497213)

[CONCLUSIONES 40](#_Toc142497214)

[REFERENCIAS 41](#_Toc142497215)

# **INTRODUCCIÓN**

En este informe, se describe el desarrollo de un software de automatización diseñado para buscar perfiles en LinkedIn. Este se desarrolló en colaboración con la empresa Forte Innovation Consulting, quieres brindaron su experiencia para implementar el sistema en la empresa Metaphorce la cual es una empresa hermana, con el objetivo de optimizar el proceso de búsqueda de talentos en LinkedIn.

Para lograr este desarrollo, se siguió el ciclo de desarrollo de software, que se divide en etapas clave: planeación, análisis, diseño, codificación, pruebas y mantenimiento. Cabe destacar que este informe se centra específicamente en las etapas de análisis, diseño y codificación.

En la etapa de planeación, se establecieron los objetivos del programa de automatización y se definieron los requerimientos necesarios para su implementación. Durante este proceso, se consideraron diferentes opciones para lograr estos objetivos, evaluando alternativas y determinando la mejor solución.

En la etapa de análisis, se llevó a cabo una investigación para buscar herramientas adecuadas que permitieran crear el programa de automatización. Se examinaron los procesos internos existentes que proporcionaban pautas sobre cómo iniciar el desarrollo, como los requerimientos técnicos utilizados por la empresa en programas similares.

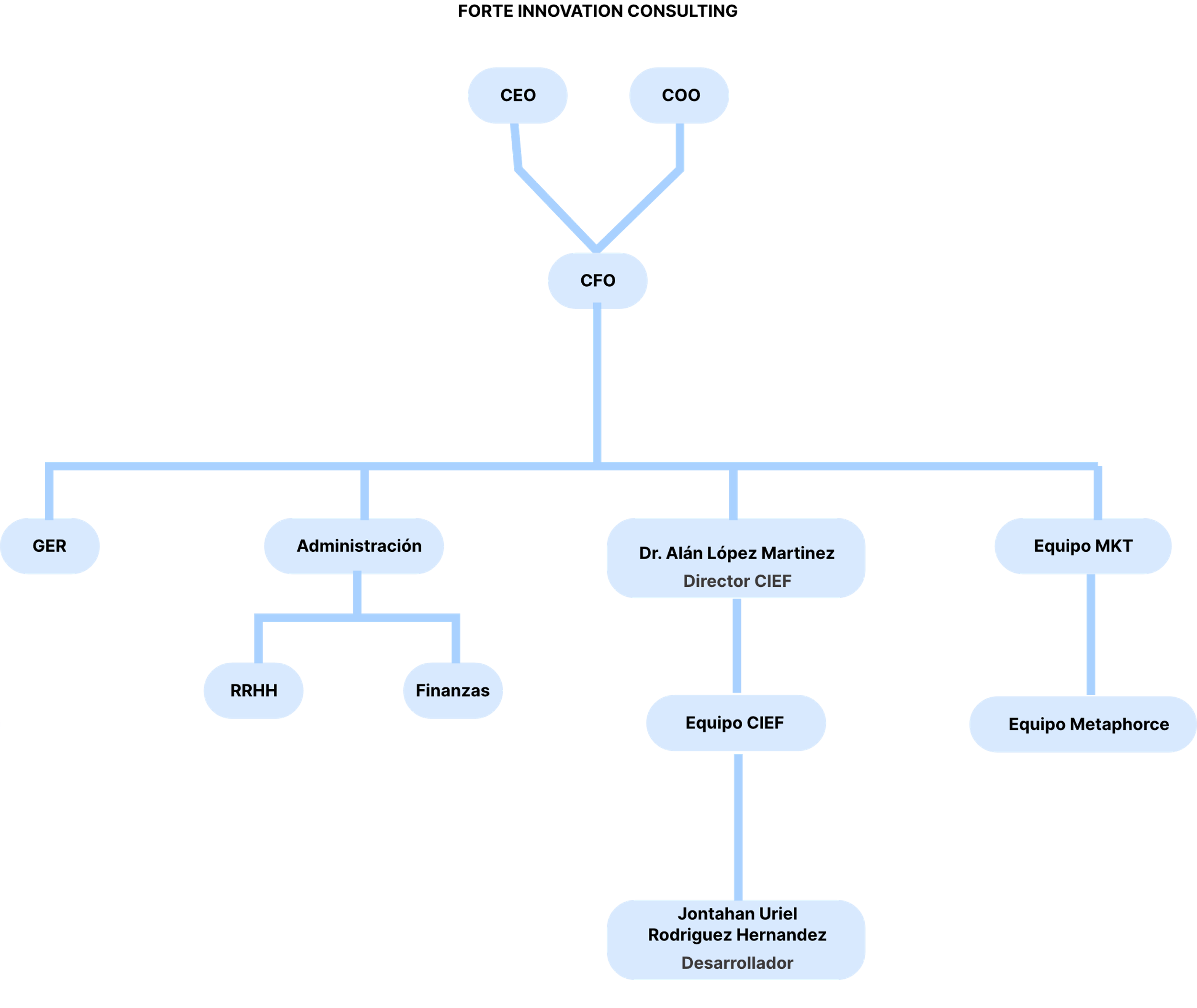
Esta etapa también se convirtió en una fase de diseño de la arquitectura del programa. Se optó por un enfoque de diseño de arquitectura modular, con el objetivo de asegurar que los componentes fueran independientes entre sí.

Por último, se procedió a la etapa de codificación, donde se escribió el código fuente del programa siguiendo las mejores prácticas y estándares de programación. Se implementaron las funcionalidades planificadas, además de agregar algunas pruebas unitarias que permitan saber si el RPA funciona correctamente al cambiar el código en un futuro.

Es importante resaltar que, aunque este informe se centra en la etapa de codificación, las fases restantes del ciclo de desarrollo de software (pruebas y mantenimiento) son igualmente críticas para asegurar el correcto funcionamiento del programa y su adaptabilidad a futuras necesidades de Metaphorce.

# **ANTECEDENTES**

Forte Innovation Consulting, empresa dedicada al desarrollo e implementación de soluciones tecnológicas, ubicada en la Plaza de La Paz, Puerto Interior #102, C.P. 36275 Silao, Gto., cuenta con más de 15 años de experiencia y ha destacado tanto en el sector público como en el privado. Como parte de su compromiso con la innovación, la empresa ha lanzado Metaphorce como uno de sus proyectos más recientes, dedicado al reclutamiento y búsqueda de talento, con la intención de mejorar la situación del sector tecnológico en la región.

El equipo de Metaphorce es un pequeño equipo de trabajadores de Forte Innovation Consulting entre ellos está el equipo de marketing y del área de CIEF (imagen 1) el cual está enfocado a tecnologías con aplicaciones avanzadas como lo es el procesamiento de lenguaje natural, blockchain, inteligencia artificial además de apoyar a los diversos equipos que hay dentro de la empresa.

**Imagen 1. Organigrama de la empresa Forte Innovation Consulting.**

Es importante decir que Metaphorce se dio cuenta que, debido al entorno actual de negocios altamente competitivos, la búsqueda y contratación de profesionales de tecnologías de la información (TI) se volvió esencial para mantener una ventaja competitiva. Sin embargo, esto se convirtió en una tarea altamente laboriosa y difícil ya que por lo regular estas tareas consumen mucho tiempo y recursos.

Por ello el objetivo del presente proyecto es lograr automatizar la búsqueda de roles en tecnologías de la información para lograr que se requiera menos esfuerzo humano en la búsqueda y análisis de candidatos para puestos de trabajo, ya que normalmente esto suele ser un trabajo manual y repetitivo.

Además, se esto ser una solución altamente innovadora para la agilización de procesos que suelen ser tardados como lo es la contratación de una persona, así como en el contexto de Metaphorce poder agregarlo a la plataforma y permita configurar la búsqueda de perfiles efectivamente.

# **PROBLEMÁTICA**

La problemática planteada se centra en el desarrollo de un prototipo de módulo de Automatización Robótica de Procesos (RPA) que se integre con la plataforma Metaphorce, con el objetivo de configurar la búsqueda de roles de Tecnologías de la Información (TI) en LinkedIn.

Además, al utilizar la plataforma de LinkedIn para obtener información, es de suma importancia cumplir con las normativas y políticas establecidas para su uso y protección de la privacidad de los datos. Al implementar robots digitales para la automatización de procesos, se debe garantizar el cumplimiento de estas normativas para mantener la integridad y confidencialidad de la información.

Por último, si no se implementa el módulo de RPA, la empresa Forte Innovation Consulting seguirá dedicando una cantidad significativa de tiempo y recursos al proceso de selección de personal. Esta situación supondrá una desventaja competitiva en un mercado tecnológico en constante evolución, donde la eficiencia y la agilidad en la contratación de talento son factores clave para el éxito empresarial.

# **JUSTIFICACIÓN**

La implementación de la RPA dentro de la plataforma Metaphorce optimizará el proceso de contratación al acelerarlo, reducir el esfuerzo humano, disminuir errores, permitir un control parametrizado, revisar un mayor volumen de datos y simplificar el contacto con los candidatos. Estos beneficios ayudarán a mejorar la eficiencia y la calidad de las contrataciones en el ámbito de los roles de TI.

Entre estos los más urgentes para solucionar es el control parametrizado, la disminución del esfuerzo humano y la revisión de grandes cantidades de datos, es importante agregar que en caso de no cumplir con la integración del módulo RPA ocasionara lo siguiente:

Atraso en las entregas del proyecto de Metaphorce: Al estar la aplicación de Metaphorce aun en etapas de desarrollo es importante que se realice esta entrega ya que en caso contrario ocasionaría un retraso en el desarrollo de nuevas funcionalidades además de esta contar con fechas de entrega al cliente.

Mayor carga de trabajo para los empleados: Sin la RPA, los empleados tendrán que invertir más tiempo y esfuerzo en realizar tareas repetitivas y laboriosas, como buscar perfiles, revisar currículums y realizar evaluaciones manuales. Esto puede generar agotamiento y disminuir la productividad en otras áreas de trabajo.

Mayor margen de error en la selección de candidatos: Al depender únicamente de la revisión manual de perfiles, existe un mayor riesgo de cometer errores en la evaluación y selección de candidatos. Esto puede resultar en la contratación de personas que no cumplen con los requisitos necesarios o que no se ajustan a la cultura y objetivos de la empresa, lo que afecta negativamente la eficiencia y el éxito de los proyectos.

Dificultad en la comunicación con los candidatos: La falta de automatización dificulta el proceso de contacto y comunicación con los candidatos seleccionados. Esto puede generar demoras en la programación de entrevistas, dificultades en la coordinación y una experiencia deficiente para los candidatos, lo que puede afectar la percepción de la empresa y su capacidad para atraer talento.

# **OBJETIVOS**

* 1. OBJETIVO GENERAL

Desarrollar un módulo de búsqueda para la localización candidatos en la plataforma de LinkedIn mediante un RPA.

* 1. OBJETIVOS METODOLÓGICOS
     1. Analizar los requerimientos funcionales para el RPA.
     2. Diseñar la arquitectura del RPA para una clara identificación de requerimientos mediante diagramas UML.
     3. Desarrollar el módulo de RPA para su utilización mediante Python.
  2. ALCANCE

Es importante destacar que el alcance del proyecto se centró solamente en las partes iniciales, lo que conllevó la configuración y automatización de las búsquedas en LinkedIn para los perfiles a través de la implementación de la RPA, así como agregarla a la plataforma de Metaphorce.

Se brindó documentación y capacitación al equipo de Metaphorce para que el módulo pudiera ser utilizado eficientemente, así como permitir que posteriormente se pudieran agregar más funcionalidades, ya que estas quedaban fuera del alcance inicial.

Como principal limitante, se debía respetar el uso de las políticas y términos de uso de LinkedIn en relación con la automatización de búsquedas e integración con los perfiles de usuario, al tener que obtener el consentimiento de los usuarios para poder utilizar sus datos, así como protegerlos.

# **METODOLOGÍA**

A continuación, se presenta la secuencia de pasos que se siguieron a lo largo de la implementación del RPA. Es importante decir que como se ve en la siguiente tabla (Tabla 1). Cronograma de actividades, tras la etapa de onboarding que consiste en entrar al equipo y lograr comprender el proyecto, justo después fue cuando se llevó a cabo las actividades específicas del análisis, diseño y programación del RPA. Estas actividades se muestran a continuación.

**Tabla 1. Cronograma de actividades**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Actividad** | **Inicio** | **Fin** | **Semana 1** | **Semana 2** | **Semana 3** | **Semana 4** | **Semana 5** | **Semana 6** | **Semana 7** | **Semana 8** | **Semana 9** | **Semana 10** | **Semana 11** | **Semana 12** | **Semana 13** |
| Onboarding | Semana 1 | Semana 3 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Revisión de historias de usuario | Semana 4 | Semana 4 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Matriz funcional | Semana 4 | Semana 4 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Diseño del RPA | Semana 5 | Semana 6 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| programación del RPA | Semana 6 | Semana 10 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Despliegue del RPA | Semana 11 | Semana 12 |  |  |  |  |  |  |  |  |  |  |  |  |  |

En resumen, para realizar la implementación del RPA se siguió un orden especifico de actividades para cumplir con el ciclo del desarrollo de software tradicional.

* 1. ANÁLISIS DE LOS REQUERIMIENTOS FUNCIONALES DEL RPA

Para comenzar un desarrollo de software de manera efectiva, es fundamental realizar un exhaustivo análisis de requerimientos. De acuerdo con Pressman (2010), este análisis "proporciona el mecanismo adecuado para comprender las expectativas del cliente, analizar las necesidades, evaluar la viabilidad, negociar una solución razonable, especificar claramente la solución, validar la especificación y administrar los requerimientos" (p. 102). Dando a entender la importancia que cumplen los requerimientos para el desarrollo de software.

A través de estos análisis se busca llegar a obtener los requerimientos funcionales para realizar la implementación de una Automatización Robótica de Procesos (RPA) el cual es una herramienta con la cual un proceso que una persona tardaría normalmente horas o incluso semanas, un RPA lo haría en minutos. Por lo cual según Kaelble (2018) “En comparación con los humanos, se ha demostrado que los robots de software completan las mismas tareas aproximadamente cinco veces más rápido. También trabajan las 24 horas del día, los 7 días de la semana, y no piden tiempo libre” (P.10). Siendo estas características extremadamente útiles para cualquier empresa.

Para obtener los requerimientos funcionales en este caso, se utilizó como base un archivo de Excel proporcionado, y se complementó con la información obtenida en varias reuniones. Sin embargo, este proceso tomó más tiempo de lo esperado debido a cambios iniciales en el equipo de Metaphorce, lo que generó retrasos en el proceso.

A pesar de estos desafíos, se hizo un esfuerzo por abordar los aspectos clave del proceso del RPA. Se optó por crear un conjunto de preguntas y respuestas basado en el conocimiento disponible, con el propósito de aclarar y definir mejor los detalles. Estas preguntas se centraron en aspectos generales en lugar de aspectos técnicos relacionados con el desarrollo del RPA.

A continuación, se presenta la Tabla 2 que contiene las preguntas formuladas. Estas preguntas tienen como objetivo brindar una visión más clara y comprensible de los aspectos esenciales del proyecto RPA.

**Tabla 2. Preguntas y respuestas sintetizadas de la entrevista.**

|  |  |
| --- | --- |
| **Pregunta** | **Respuestas sintetizadas** |
| ¿Cuáles son las funcionalidades  clave que debe tener el módulo de búsqueda? | Poder buscar perfiles en LinkedIn, así como  obtener sus datos de contacto |
| ¿Qué criterios de búsqueda son  relevantes para encontrar candidatos? | Nos interesa poder buscar por puesto de trabajo por ejemplo si es una persona especializada en cierta tecnología, así como su locación |
| ¿Es necesario realizar búsquedas  avanzadas con filtros específicos? | Poder configurar el número de perfiles que podemos ver así como la opción de descartar si no cuentas con datos de contacto |
| ¿Se requiere la posibilidad de  guardar búsquedas o resultados? | si se debieran de guardar las búsquedas para evitar tener que volver a buscar |
| ¿Qué tipo de información se desea  obtener de los perfiles de los candidatos? | Saber sus datos personales, datos de contacto, en tecnologías se especializa, si cuenta con carrera o aprendió por cursos, si cuenta con certificaciones y los idiomas que habla, así como su nivel en estos |
| ¿Es necesario integrar el módulo de  búsqueda con otras funcionalidades  o sistemas existentes? | Si necesitamos que este integrado con la plataforma de Metaphorce para poder realizar las búsquedas desde ahí, así como un dashboard para ver la información |
| ¿Cuáles son los plazos y  prioridades del proyecto? | Debería de estar listo para la entrega del MVP 5 cuya fecha de entrega es el 28 de julio |

Por otro lado, también se proporcionó los requerimientos técnicos específicos para el proyecto. Entre estos requerimientos, se determinó el lenguaje de programación a utilizar, optando por Python debido a la familiaridad y comodidad del equipo de Metaphorce al trabajar con este lenguaje. Además, Python es ampliamente utilizado para la automatización y como lenguaje de programación para servidores.

Además, se hizo hincapié en las posibles limitaciones en la búsqueda de librerías que podrían agilizar el desarrollo del RPA. Dado que no era necesario crear funcionalidades desde cero, por eso se exploraron librerías especializadas para facilitar este proceso. El objetivo era identificar cualquier restricción o dificultad en la disponibilidad de librerías específicas y considerar alternativas viables para optimizar el desarrollo del RPA.

Por otro lado, se determinó la plataforma de la nube en la cual se alojará el RPA y se decidió qué tipo de arquitectura se utilizaría para la comunicación con la plataforma de Metaphorce y la forma de lograr su correcta integración.

Una vez finalizada esto, se dio inicio a un minucioso proceso de análisis para obtener la lista definitiva de requerimientos funcionales del RPA. Durante este proceso, se identificaron los aspectos cruciales del funcionamiento del RPA y se exploraron diferentes enfoques para su implementación.

En el análisis, se tuvieron en cuenta las limitaciones de la plataforma de Metaphorce, asegurándose de que los requerimientos fueran factibles y realistas dentro de este entorno. Se puso énfasis en la simplicidad y claridad de los requerimientos para evitar cualquier tipo de malentendidos durante el desarrollo y la posterior implementación del RPA.

El resultado de este análisis fue una lista concisa de requerimientos funcionales que capturaba de manera precisa las necesidades del proyecto y los objetivos comerciales.

En resumen, podemos concluir que la combinación de la información proporcionada en reuniones y otros canales resultó ser un enfoque sólido para capturar los requerimientos funcionales. Sin embargo, es esencial tener en mente que estos requerimientos podrían no abarcar aspectos técnicos detallados en ciertos casos. Por lo tanto, se aconseja contar con la guía de un experto experimentado, quien pueda enriquecer los requisitos generales con elementos técnicos esenciales.

* 1. DISEÑO DE LA ARQUITECTURA DEL RPA PARA UNA CLARA IDENTIFICACIÓN DE REQUERIMIENTOS MEDIANTE DIAGRAMAS UML

Por lo general, al comenzar un desarrollo de software desde cero, es común diseñar el sistema después de obtener los requerimientos. En este proceso, los diagramas desempeñan un papel crucial, ya que sirven como base para implementar el programa a través de código.

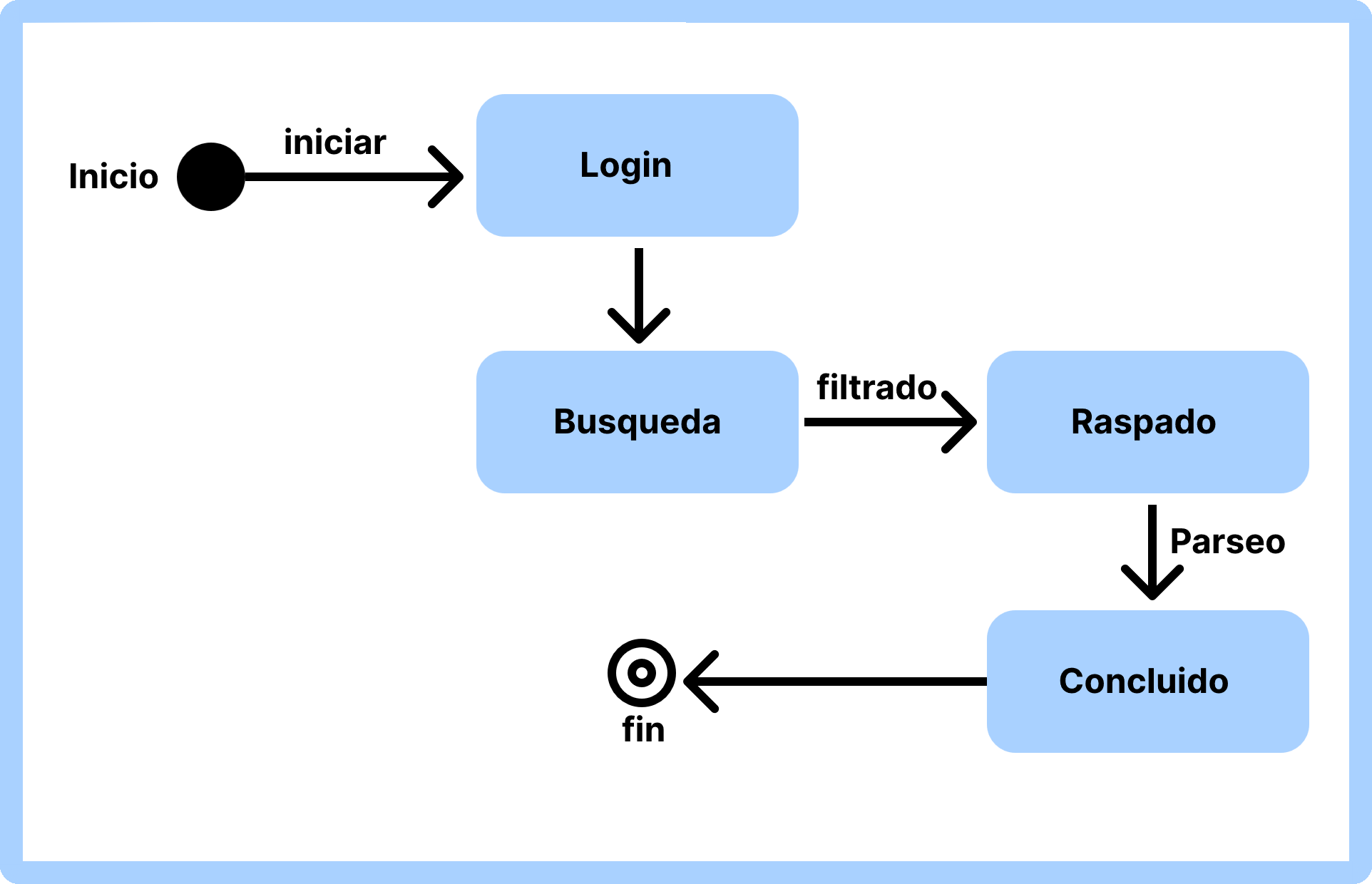
Uno de los conjuntos de gráficos más utilizados en el diseño de software es el Lenguaje Unificado de Modelado (UML), el cual proporciona estándares y notaciones para realizar estos diagramas.

Cabe mencionar que de acuerdo con Rumbaugh et al (2007) “El último objetivo de UML era ser tan simple como fuera posible, pero manteniendo la capacidad de modelar toda la gama de sistemas que se necesita construir” (P.9). Lo cual para un desarrollo lo vuelve muy util y versátil para documentar y facilitar lo que se deber programar.

Dentro de la elaboración de estos el diagrama de estado, también conocido como diagrama de máquina de estados, desempeñó un papel crucial en el modelado y visualización del comportamiento del sistema durante la ejecución del programa.

Su objetivo principal fue comprender y analizar cómo el sistema respondía a diversos eventos y cómo se desarrollaba el flujo de trabajo durante la ejecución de las tareas automatizadas. Este diagrama describe el comportamiento dinámico de los objetos a lo largo del tiempo, representando los ciclos de vida de manera precisa y detallada (Rumbaugh et al., 2007, p. 75). Por ello es por lo que se decidió hacer uso de este diagrama.

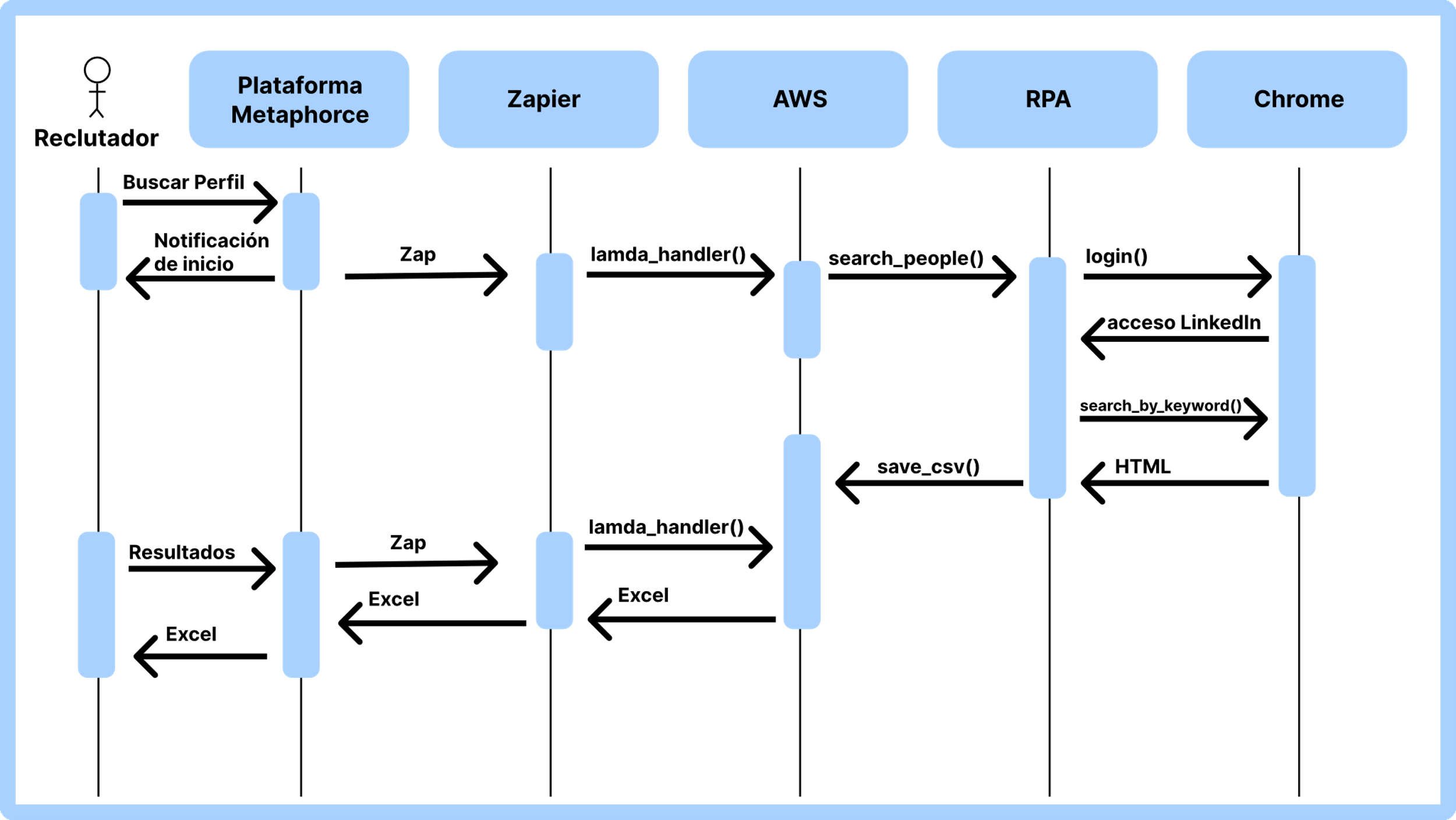
En la siguiente imagen (imagen 2). Se ilustra el proceso de inicio del RPA a través de una configuración inicial. A continuación, el sistema realiza un inicio de sesión en LinkedIn y, mediante una búsqueda y filtrado de información, realiza el raspado de datos para encontrar personas. Finalmente, concluye depositando la información obtenida en la plataforma de Metaphorce.



**Imagen 2. Diagrama de estado del RPA.**

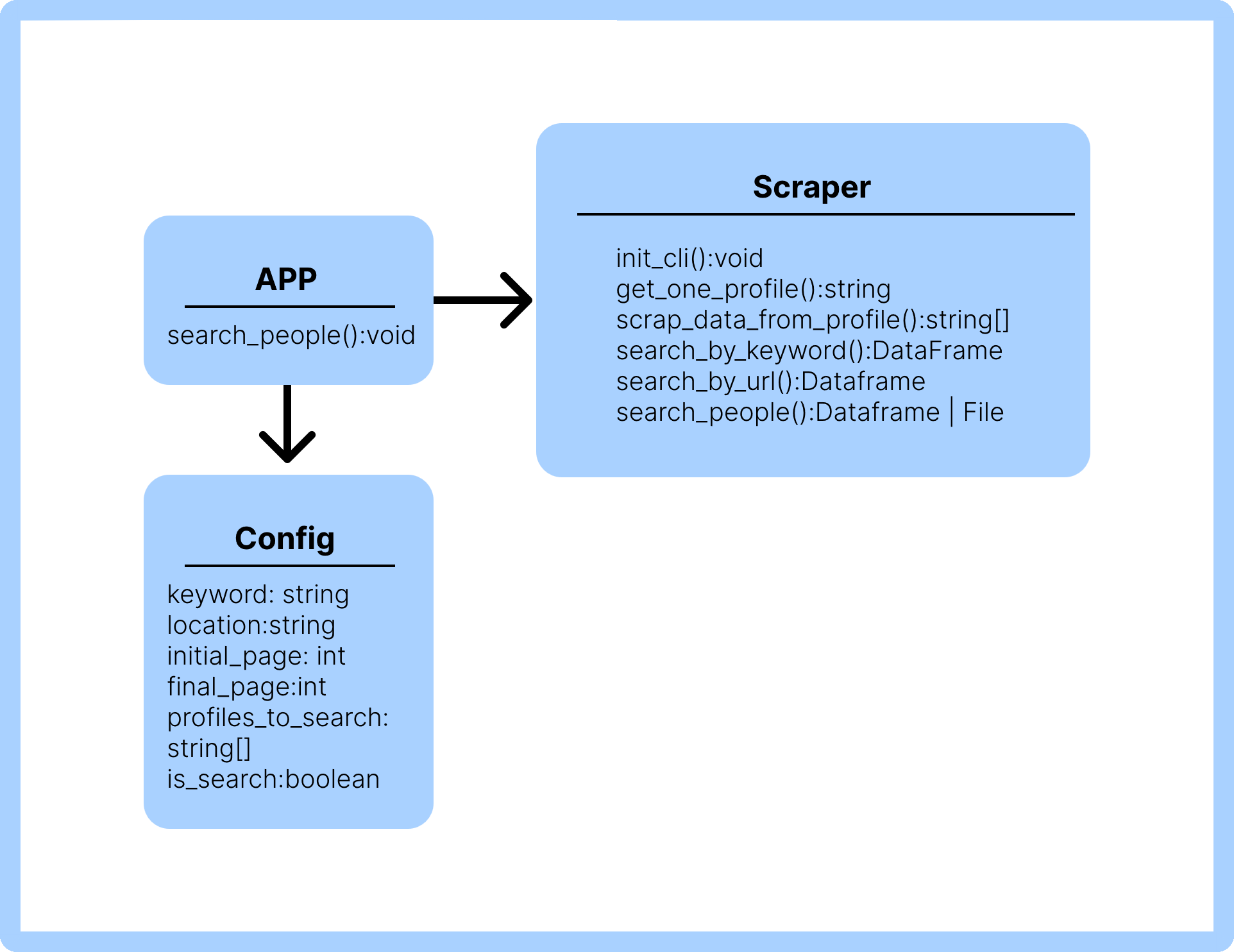
Por otro lado, el diagrama de secuencia permitió trazar los pasos y etapas clave de la interacción, mostrando claramente cómo el usuario interactuaba con el sistema en cada punto y cómo se generaban las respuestas correspondientes. Proporcionó una visión global y estructurada de la secuencia de interacciones, lo que facilitó la identificación de puntos críticos de integración y posibles mejoras en la experiencia del usuario.

Adicionalmente, en la imagen (imagen 3) se ilustra la interacción específica del reclutador en la plataforma. En este caso, el reclutador realiza una búsqueda de perfil que es enviada a una función Lambda de AWS mediante zapier, la cual inicializa un RPA al finalizar, el RPA devuelve la información encontrada al reclutador. Este diagrama ofrece una representación visual clara de la secuencia de acciones y procesos involucrados en la interacción del reclutador con la plataforma.

****

**Imagen 3. Diagrama de secuencia del RPA.**

Por último, se creó un diagrama de clases como referencia al iniciar el desarrollo, el cual se muestra en la siguiente imagen (Imagen 4). En este diagrama se representaron tres clases principales: "APP", que actúa como el punto de entrada de la aplicación. "Config", que modela la configuración del RPA y "Scraper", que contiene la lógica del RPA.



**Imagen 4. Diagrama de clases del RPA.**

En conclusión, los diagramas son parte esencial del diseño de software, ofreciendo una visión panorámica y organizada de las clases y sus interacciones en un proyecto. Al utilizar estos diagramas antes de la implementación, se establece una base sólida para el desarrollo, lo que facilita la comprensión entre los miembros del equipo. También sirven como una guía visual que ayuda a prever posibles problemas o incoherencias en el diseño ya que promueven la claridad, coherencia y eficiencia en todo el ciclo de vida del proyecto, permitiendo correcciones tempranas y evitando costosos retrabajos más adelante.

* 1. DESARROLLO DEL MÓDULO DE RPA PARA SU UTILIZACIÓN MEDIANTE PYTHON

Después de completar las etapas de análisis y diseño llego el momento de pasar a la etapa de desarrollo. Sin Embargo, en esta etapa se hicieron muchas mini tareas mostradas en siguiente imagen (imagen 5) y son las que serán explicadas a continuación.



**Imagen 5. tareas hechas para hacer el desarrollo.**

Lo primero que se hizo fue configurar el entorno de desarrollo usando Inicio de desarrollo de Windows (imagen 6) la cual permite instalar los programas y lenguajes necesario para poder empezar a programar en este caso se instaló Python en su versión 3.11, git para versionamiento de código y visual estudio code como IDE de desarrollo.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

**Imagen 6. Instalación en Inicio de desarrollo (Microsoft, 2023).**

Una vez se completó la configuración inicial para comenzar la programación del RPA, se llevó a cabo el siguiente proceso:

* Creación de la carpeta y configuración del repositorio de git: Se creo una carpeta específica para alojar el código del RPA. Además, se inicializo un repositorio de Git para facilitar el control de versiones del código. Git es una herramienta que “Podría entenderse como un sistema para crear copias de seguridad de nuestro código, y en gran parte así es, pero su uso también se extiende a cómo lidiamos con las actualizaciones de ese código y lo sincronizamos entre diferentes personas” (Duran,2023, P.12). El cual puede verse en la imagen (imagen 7) donde se ve el primer estado del repositorio.

Captura de pantalla de un celular

Descripción generada automáticamente

**Imagen 7. Repositorio de código del RPA (GitHub, 2023).**

* Creación del archivo principal. Se creo un archivo inicial que serviría como punto de inicio del RPA. Este archivo contendría la estructura básica y las configuraciones iniciales para iniciar el desarrollo del RPA. En la imagen (imagen 8) donde se puede mirar esta primera estructura.

Texto

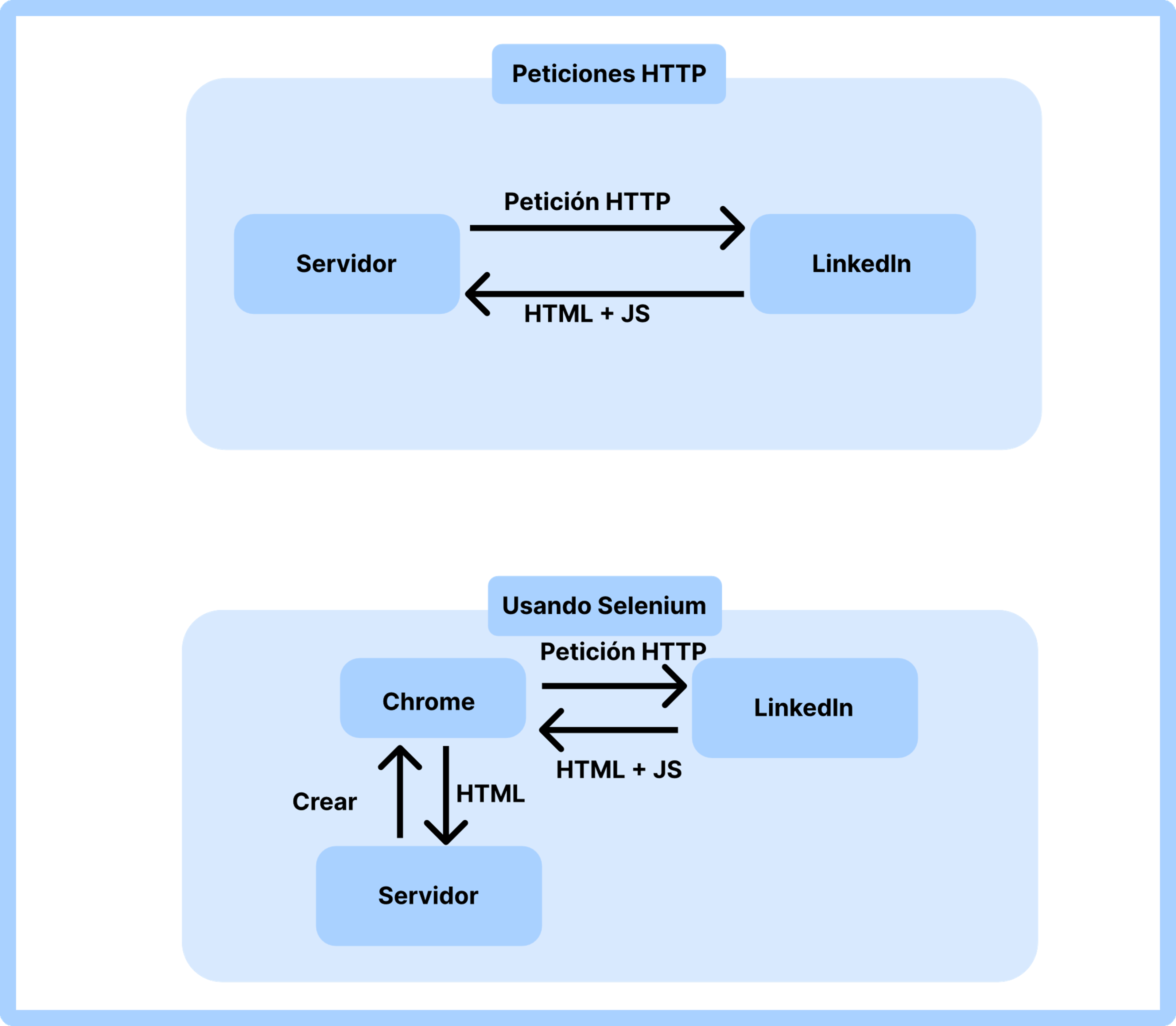
Descripción generada automáticamente

**Imagen 8. Estructura básico del RPA (GitHub,2023).**

* Instalación de las bibliotecas de código necesarias. Se procedió a instalar las librerías que proporcionaron funcionalidades adicionales y herramientas especializadas con la automatización. Para instalación de librerías se utilizó Python Package Index (pip), un gestor de paquetes de Python que permite instalar, actualizar y administrar bibliotecas de terceros.
* Revisión de la página de LinkedIn. Durante la revisión de la página de LinkedIn y tras un análisis se determinó que el contenido de los perfiles se generaba mediante renderizado en el cliente (CSR, por sus siglas en inglés). Esto significa que el código HTML de la página no se envía directamente desde el servidor, sino que se genera y completa en el navegador del cliente. HTML (HyperText Markup Language) es el lenguaje de marcado estándar utilizado para estructurar y presentar el contenido en la web.

Debido a una característica particular, el RPA enfrento una limitación en el acceso directo al código fuente de LinkedIn mediante solicitudes HTTP regulares. En este caso, es importante destacar que HTTP (Protocolo de Transferencia de Hipertexto) es el protocolo de comunicación utilizado para transferir datos en la web.

En cambio, se identificó la necesidad de crear una instancia de un navegador usando la biblioteca “Selenium” (imagen 9) para interactuar con LinkedIn y extraer la información requerida. Al adoptar esta estrategia, el RPA sería capaz de simular la interacción de un usuario real con LinkedIn, permitiendo acceder a los perfiles y extraer la información deseada de manera precisa y eficiente.



**Imagen 9. Diferencia de usar peticiones HTTP y Selenium.**

Una vez completados los pasos previos, se procedió a iniciar la programación del RPA. Siguiendo los requerimientos establecidos, se comenzó por declarar una clase para el RPA, que recibiría un diccionario como parámetro para su configuración y uso adecuado.

Posteriormente, se instaló la librería "Selenium". Esta librería fue fundamental para poder acceder al código de la página de LinkedIn, superando los desafíos asociados con el CSR. Con Selenium, se logró cargar la página en el navegador automatizado y obtener su código fuente sin ningún problema.

Una vez obtenido el código de la página, se utilizó la biblioteca "BeautifulSoup". Esta biblioteca permitió buscar elementos específicos dentro del HTML, así como se puede observar en la imagen (imagen 10). La información se encuentra mediante búsquedas basadas en etiquetas, clases o identificadores de elementos, la cual se puede extraer fácilmente con BeautifulSoup.

Texto

Descripción generada automáticamente

**Imagen 10. HTML del perfil de pruebas (Firefox,2023).**

La combinación de Selenium y BeautifulSoup resultó ser una solución efectiva para interactuar con el navegador y realizar la extracción de datos necesaria para el RPA. Selenium permitió la interacción automatizada con la página, mientras que BeautifulSoup facilitó el análisis y la extracción precisa de elementos HTML.

En el proceso de extracción de datos, se priorizó la recopilación de información relevante que permitiera a los profesionales de reclutamiento evaluar la capacidad de los candidatos para un puesto determinado. Para lograr esto, se identificaron varios datos clave que debían ser extraídos de los perfiles de LinkedIn.

Los cuales fueron nombre del candidato, la imagen del perfil, si contaba con una descripción personal, los idiomas que habla tanto como su nivel de dominio en cada uno, la experiencia laboral previa en roles similares, la existencia de certificaciones relevantes y sus antecedentes educativos.

Una vez se extrajeron los datos correctamente, el siguiente paso consistió en enviar una configuración de búsqueda al RPA. Esta configuración incluía información como el puesto deseado, la ubicación donde se realizaría la búsqueda y las URL específicas a considerar.

Para lograr esto, se desarrolló una página web provisional utilizando un marco de trabajo llamado "Vue". El cual de acuerdo con Hanchett y Listwon “se puede incorporar a una página web existente para tareas simples o que se puede utilizar como base para una aplicación web a gran escala” (2018, P.16). Lo cual lo hace perfecto para crear esta página web.

En la imagen adjunta (imagen 11), se puede observar la sencillez de la página web, con campos para ingresar el puesto, la ubicación y el número de páginas de información que el RPA deberá obtener.

Interfaz de usuario gráfica, Aplicación, Teams

Descripción generada automáticamente

**Imagen 11. Pagina provisional de configuración del RPA (Firefox,2023).**

Para establecer la conexión entre el RPA y la página web, se utilizó Flask para crear un endpoint. En este contexto, un endpoint es un punto de acceso o una URL específica que actúa como un punto de comunicación entre dos sistemas.

Mientras, Flask de acuerdo con Grinberg (2014) “es un marco de trabajo pequeño según la mayoría de los estándares, lo suficientemente pequeño como para ser llamado un micro-framework” (P.3). Lo cual lo hace fácil de utilizar para crear los endpoint con los cuales se podrá comunicar el RPA.

Cuando el RPA necesita interactuar con la página web, envía una solicitud al endpoint creado utilizando métodos HTTP como GET o POST. Siendo GET para obtener la información al servidor y POST para enviar la informacion al servidor, de esta manera, se establece una conexión y comunicación bidireccional entre el RPA y la página web a través del endpoint definido en Flask.

Para este caso, la página web envía la configuración de búsqueda al RPA utilizando el formato JSON (JavaScript Object Notation). Ya que de acuerdo con Haverbeke “Es ampliamente utilizado como un formato de almacenamiento y comunicación de datos en la Web, incluso en otros lenguajes diferentes a JavaScript” (2018, P.81). Lo cual lo hace que sea perfecto para comunicar la web con el RPA.

Como resultado, el reclutador solo necesita realizar una búsqueda en la página web, lo que generará una petición HTTP de tipo POST con la configuración en formato JSON. Y dicha petición se enviará al servidor mediante el endpoint creado con Flask, que iniciará la búsqueda de datos del RPA.

Además, se agregaron test unitarios al código para observar la ejecucion del sistema con la finalidad de validar el comportamiento previsto o en caso contrario encontrar fallas (Irrazábal y Mascheroni, 2022, P.1). Con lo cual se mejoraría el proceso de entrar errores al momento de cambiar cosas dentro del código fuente del RPA.

Estas pruebas unitarias consistieron en probar que el código verifique ciertos datos al momento de buscar los perfiles dentro de LinkedIn, validar de no dar valores errados, así como validar que al final regrese un archivo con los datos de los perfiles el cual se puede observar en la siguiente imagen (imagen 12).

Texto

Descripción generada automáticamente

**Imagen 12. Porción del código de las pruebas unitarias (VScode,2023).**

Como último paso, se decidió utilizar archivos markdown para crear documentación técnica ya que “Markdown es una opción natural para la documentación técnica. Empresas como GitHub están adoptando cada vez más Markdown para su documentación” (Cone, 2020, P.10). Además, su popularidad, Markdown es fácil de implementar y leer. Lo que agiliza la tarea de documentar el progreso del proyecto.

Esta documentación cumple con la idea de crear un mapa que guía a los desarrolladores en el proceso de implementación y mantenimiento del RPA. Además, sirve como un recurso valioso para futuras actualizaciones, mejoras y la resolución de problemas que puedan surgir.

Para concluir, el desarrollo del RPA implicó enfrentar diversos desafíos que se superaron mediante un proceso de prueba y error. Uno de los aspectos clave en este proceso fue la extracción de la información necesaria. Además, se llevó a cabo una labor significativa al redactar la documentación correspondiente.

Sin embargo, al contar con requerimientos bien definidos y una arquitectura de proyecto clara facilitó en gran medida el desarrollo del RPA. Esto permitió llevar a cabo las tareas de manera más eficiente y efectiva, asegurando el cumplimiento de los objetivos establecidos.

# **RESULTADOS**

Después de completar la metodología, es fundamental realizar una descripción detallada de los resultados obtenidos. Durante esta etapa, se presentarán de manera clara y precisa los resultados alcanzados en cada uno de los puntos descritos en la metodología.

* 1. ANÁLISIS DE LOS REQUERIMIENTOS FUNCIONALES DEL RPA

Como se explicó anteriormente después de pasar por el proceso de entrevista y analizar los requerimientos técnicos, se crearon los requerimientos funcionales necesarios para implementar el RPA mediante un exhaustivo análisis de las necesidades y objetivos de la organización, así como de las tareas y procesos específicos que se buscaban automatizar.

Por ello con el propósito de describir las necesidades del cliente, establecer una base para el diseño de software y definir un conjunto de requerimientos validables, se siguen los objetivos principales mencionados por Pressman (2010). Estos objetivos se presentan en la siguiente tabla (tabla 2), que muestra el número, nombre y descripción de cada requerimiento.

**Tabla 2. Tabla de requerimientos del RPA.**

|  |  |  |
| --- | --- | --- |
| Requerimiento # | Nombre del requerimiento | Descripción del requerimiento |
| CU-001 | Uso de LinkedIn | Se requiere que la extracción de datos se haga mediante la plataforma de LinkedIn |
| CU-002 | Login en LinkedIn | EL RPA deberá permitir el iniciar sesión en una cuenta de LinkedIn para poder acceder a la información. |
| CU-003 | Configuración de búsqueda | Se debe permitir hacer una configuración para realizar la búsqueda de los perfiles.  Los campos de configuración deben de ser: Puesto de trabajo que se está buscando, locación donde se requiere buscar (si es remoto puede ser cualquier lugar), numero de perfiles que se desean buscar, número de páginas en las que se realizara la búsqueda. |
| CU-004 | Extracción de datos | Una vez el RPA pueda iniciar sesión en LinkedIn el RPA debe obtener la siguiente información de cada perfil: URL del perfil, nombre, imagen, descripción, empresa actual, descripción, idiomas, educación, experiencia, conocimiento y actitudes, así como licencias y certificaciones |
| CU-005 | Almacenamiento de la información extraída | De manera provisional la informacion extraída se guardará en formato, CSV.  Con posibilidad de guardarse en el almacenamiento S3 de AWS |
| CU-006 | Creación de reportes del candidató | La aplicación debería permitir realizar la búsqueda por uno  o varios candidatos a través de su URL del perfil y  se podrá ver un reporte de o los candidatos en formato de Excel |
| CU-007 | Despliegue del RPA | El RPA debe estar desplegado en la plataforma de  Metaphorce usando lambdas para la comunicación |
| CU-008 | Comunicación entre WP-Forms y AWS Lambda | Para permitir la comunicación entre el WP-Forms y las lambdas de AWS  se debe usar la plataforma de zapier |

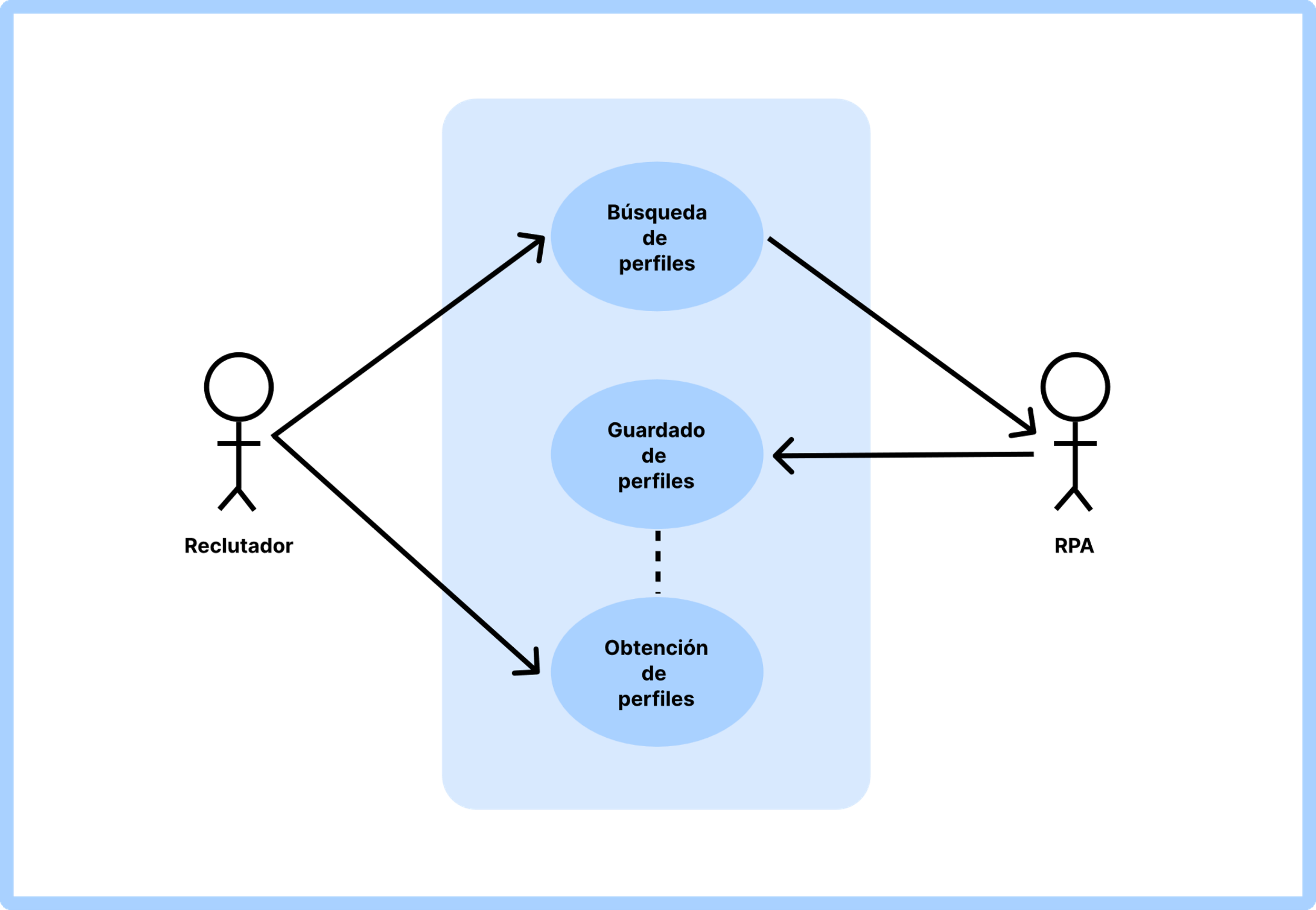
Siendo así estos los requerimientos que serían la estructura central de lo que se empezaría a diseñar y codificar en las siguientes etapas del desarrollo de este RPA.

* 1. DISEÑO DE LA ARQUITECTURA DEL RPA PARA UNA CLARA IDENTIFICACIÓN DE REQUERIMIENTOS MEDIANTE DIAGRAMAS UML

Durante el proceso de diseño de la arquitectura para crear el RPA, se llevaron a cabo distintos análisis para determinar cómo este sistema debería funcionar, qué procesos abordaría y cómo se comunicaría con otros sistemas.

Uno de estos análisis fue realizar un diagrama de casos de usos, el cual sirve para visualizar las acciones que distintos actores y el sistema pueden realizar, esto sería util para hacer una estimación de los recursos y poder garantizar disponibilidad en él sistema.

A continuación, en la imagen (imagen 13) se muestra el diagrama resultante donde hay 3 posibles casos de uso, y dos actores, en los cuales el reclutador puede hacer una búsqueda de perfiles y obtener los perfiles gracias a que el RPA realiza la búsqueda y guarda la información.



**Imagen 13. Casos de uso del RPA**

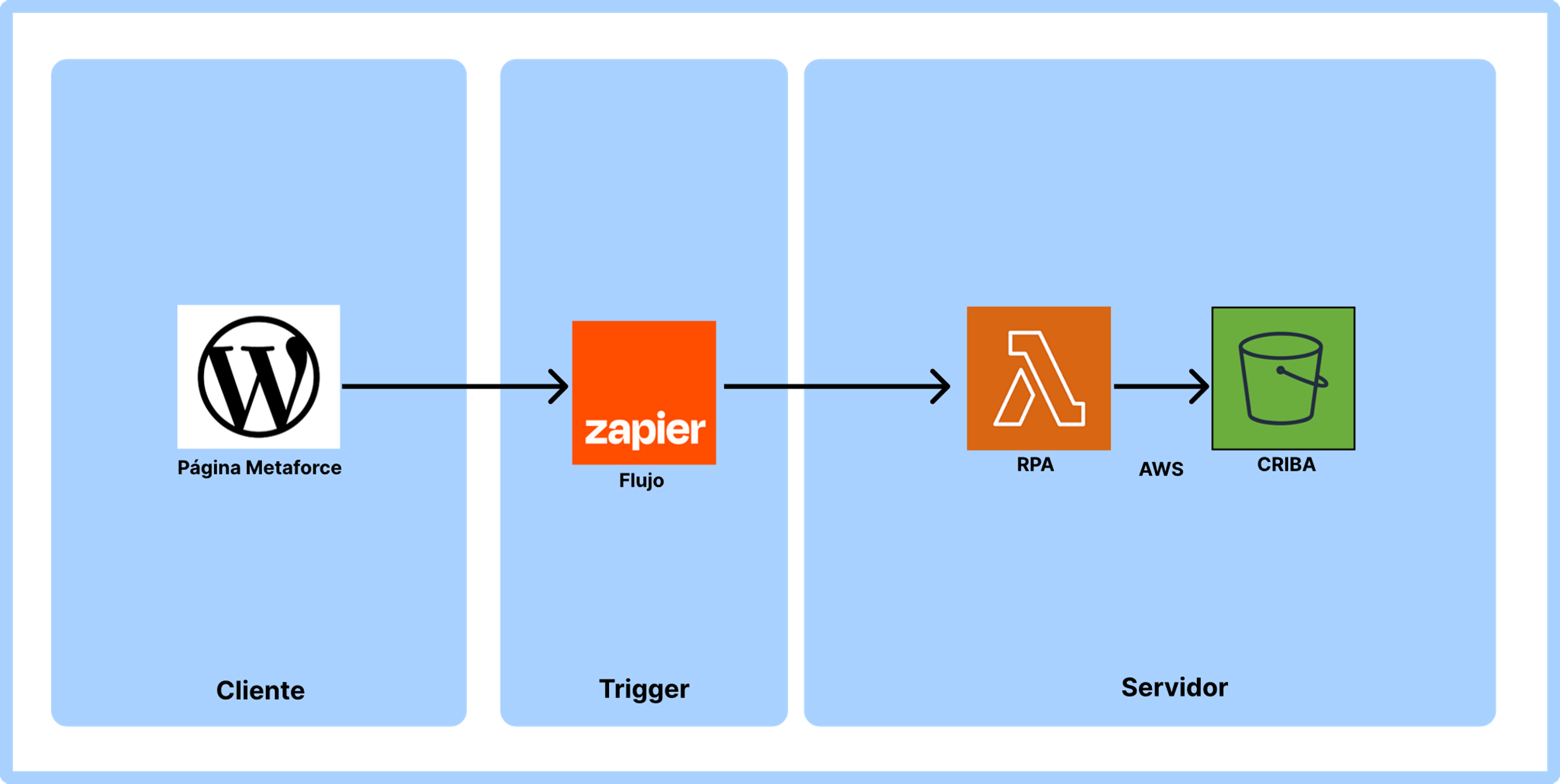
Luego del análisis se decidió que el RPA estaría alojado en AWS Lambdas, que es un servicio de computación sin servidor proporcionado por Amazon Web Services. Las Lambdas permiten una manera de ejecutar código de manera eficiente y escalable, proporcionando una infraestructura confiable para la implementación del RPA.

Además, el almacenamiento de la informacion encontrada por el RPA se almacenaría en AWS S3, el cual es un servicio de AWS para poder almacenar grandes cantidades de datos.

En cuanto a la comunicación, se optó por utilizar Zapier, una plataforma de integración que facilita la conexión y la automatización de diversas aplicaciones y servicios. Zapier actúa como un puente entre diferentes sistemas, permitiendo que el RPA interactúe con otras herramientas y aplique acciones en función de los eventos y los datos recibidos. Esto posibilita una integración fluida y una amplia gama de posibilidades para el RPA.

Por otro lado, la página web donde se realizará la configuración del RPA estará en WordPress. Esto se debe a las limitaciones del MVP (Producto Mínimo Viable, por sus siglas en inglés), donde se busca desarrollar una versión inicial del producto que cumpla con los requisitos mínimos. WordPress es un sistema que permite gestionar contenido y realizar páginas web. Sin embargo, es importante destacar que esta elección no limita futuras expansiones o integraciones con otros sistemas, ya que el diseño de la arquitectura permite la flexibilidad necesaria para adaptarse a diferentes entornos y tecnologías.

En la siguiente imagen (imagen 14) se puede observar la representación final de como de la arquitectura, con lo mencionado anteriormente el cliente siendo una página de WordPress, la utilización de Zapier como puente entre el cliente y AWS, donde se ejecutará el RPA y guardará los datos en un S3.



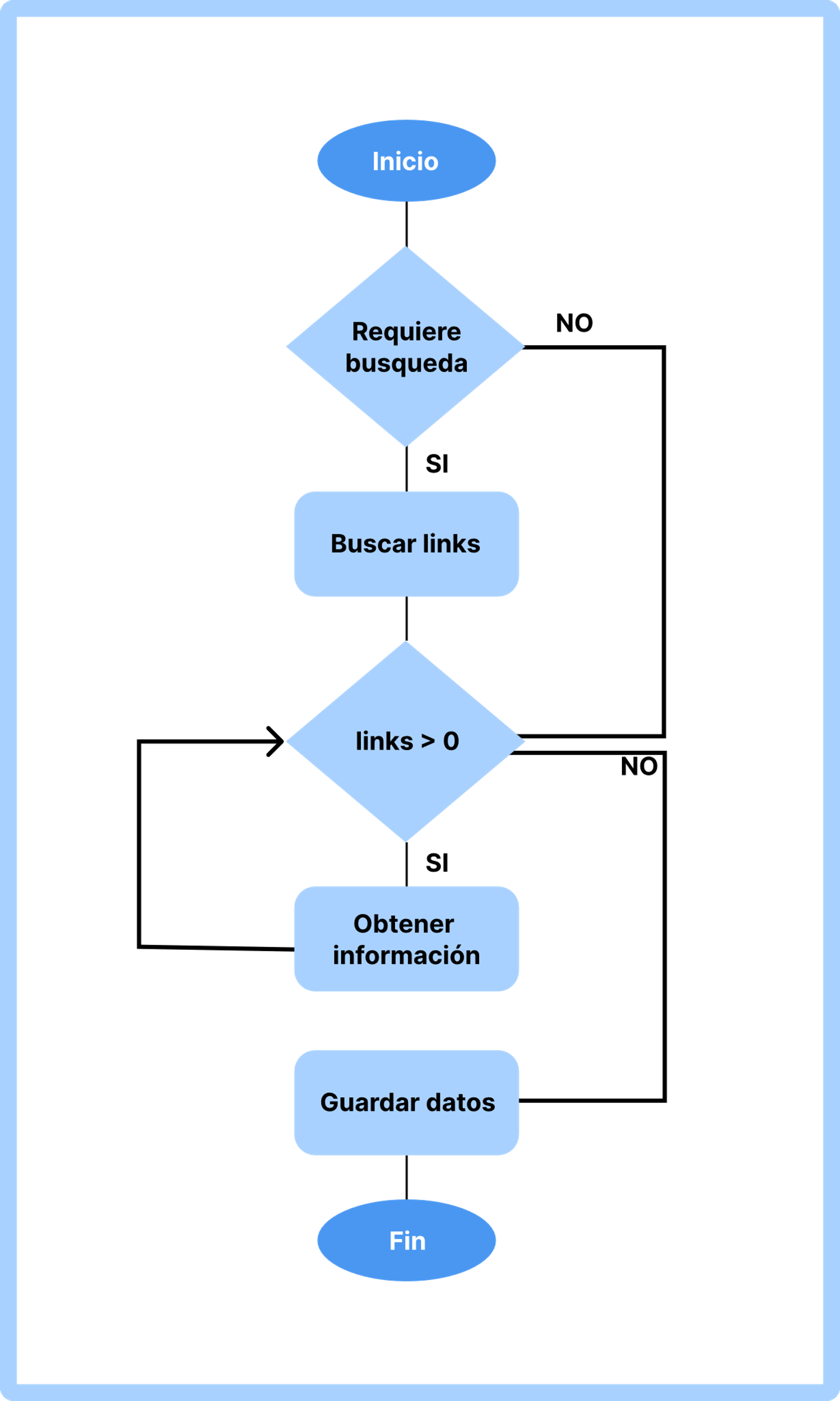
**Imagen 14. Arquitectura del RPA**

Al crear un RPA, el objetivo principal es garantizar un sistema confiable y escalable que proporcione un control completo sobre los procesos. Además, se busca facilitar el inicio, la detención y la pausa de estos. Rumbaugh et al (2018, p.12). Debido a esto y a los requerimientos se pudo llegar a la conclusión de que los objetivos para realizar el RPA serian:

1. Automatización fácil y eficiente: El objetivo principal del RPA es simplificar y agilizar los procesos mediante la automatización. Además de optimizar la eficiencia, el RPA también debe permitir búsquedas fáciles y rápidas. Esto implica la capacidad de realizar consultas y búsquedas avanzadas, proporcionando una forma fácil y eficiente de obtener los datos necesarios.
2. Flexibilidad y modularidad: La arquitectura del RPA debe diseñarse de manera que sea flexible y modular, lo que significa que los componentes individuales pueden ser modificados o reemplazados fácilmente sin afectar el funcionamiento general del sistema.
3. Desarrollo ágil y adaptable: El código del RPA debe ser diseñado teniendo en cuenta la adaptabilidad a los cambios en futuras iteraciones. Esto implica una estructura modular y bien organizada, utilizando patrones de diseño que faciliten la extensibilidad y la modificación del código. Además, es importante aplicar buenas prácticas de desarrollo, como la documentación adecuada y el uso de estándares de codificación.

Finalmente, se elaboró un diagrama de flujo que serviría como guía para el RPA, mostrando el flujo de trabajo para llevar a cabo las búsquedas de manera comprensible para cuando este se desarrolló. Este diagrama permitirá visualizar de manera clara y detallada cómo el RPA debe realizar las tareas de búsqueda, facilitando su entendimiento y uso por parte del personal del área de reclutamiento.

Esta representación (imagen 15) se muestra el diagrama, en el cual primero se requiere saber si es una búsqueda por o si ya se cuentan con perfiles específicos, en caso de que se requiera una búsqueda el RPA, buscara links de posibles candidatos a los cuales en ambos casos al final les extraerá la informacion y se guardaran los datos.



**Imagen 15. Flujo de búsqueda del RPA**

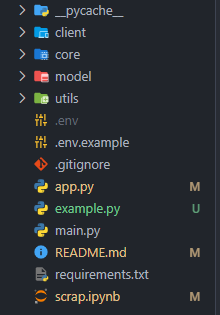
En conclusión, la arquitectura diseñada para el RPA se centró en lograr una automatización eficiente, así como en garantizar la flexibilidad y adaptabilidad del sistema. Gracias a la cuidadosa selección de tecnologías y enfoques estratégicos, se pudieron alcanzar con éxito estos objetivos fundamentales, proporcionando una sólida base para el desarrollo y la evolución continua del sistema. Es importante destacar que se le dio una gran importancia a la utilización de diagramas de UML para realizar un análisis exhaustivo y fundamentado que respaldara esta arquitectura.

* 1. DESARROLLO DEL MÓDULO DE RPA PARA SU UTILIZACIÓN MEDIANTE PYTHON

Durante el proceso de desarrollo del módulo de RPA, se logró diseñar un código que permite la automatización fácil y eficiente de búsquedas de información de personas en LinkedIn. Para lograr esto, se ha diseñado un código que permite una fácil adaptación y cambio de componentes.

Por ejemplo, el código cuenta con el uso de un patrón de diseño llamado “Adapter”, que, aunque requiere un poco más de esfuerzo en su escritura, proporciona una mayor flexibilidad ante futuros cambios (Gamma et al, 2003, p. 139). Así, se ha cuidado meticulosamente el código para garantizar su escalabilidad y modificabilidad.

Un aspecto clave para lograr esta flexibilidad se encuentra en la organización de las carpetas en el desarrollo del módulo de RPA la cual se muestra en la imagen (imagen 16). Se ha buscado separar y estructurar de manera lógica los diferentes aspectos del sistema. Por ejemplo, se ha creado una carpeta "core" para la lógica principal, una carpeta "client" para el código del cliente, una carpeta "model" para las configuraciones y una carpeta "utils" para las utilidades y funciones auxiliares. Esta estructura fomenta la modularidad, la legibilidad y la facilidad de mantenimiento del código del RPA.



**Imagen 16. Estructura de carpetas del RPA (VScode, 2023).**

Dentro del punto de entrada del RPA, se utiliza una aplicación desarrollada utilizando el marco de trabajo Flask, como se mencionó anteriormente. Esto se puede observar en la imagen proporcionada (imagen 17), donde se muestra la implementación de dicha aplicación.

Texto

Descripción generada automáticamente

**Imagen 17. Código del punto de entrada del RPA (VScode, 2023).**

El propósito de esta implementación es permitir que el cliente envíe las configuraciones necesarias para el funcionamiento del RPA. Una vez que estas configuraciones son recibidas por la aplicación, se inicia el proceso de ejecución del RPA.

La aplicación Flask proporciona una interfaz web a través de rutas definidas, como “/api/v2/people”, que son accesibles mediante solicitudes POST. Estas rutas están diseñadas para recibir los parámetros de configuración enviados por el cliente, como palabras clave, número de páginas, título, ubicación y otros detalles relevantes para el proceso de automatización.

Una vez que se reciben las configuraciones, se crean instancias de las clases correspondientes, como “ConfigScrap”, para representar y almacenar estas configuraciones. Luego, se invocan las funciones correspondientes, como “scrappProfiles” o “search\_people”, que se encargan de ejecutar el RPA y realizar las tareas de extracción de datos o búsqueda de perfiles según las configuraciones proporcionadas.

Una vez finalizado el proceso de búsqueda, el RPA devolverá un archivo con un formato similar al mostrado en la imagen (imagen 18). Este archivo estará en formato Excel, como parte de los requisitos. En él, se proporcionará la información relevante de cada perfil, necesaria para que el equipo de reclutamiento pueda determinar si la persona es apta para el puesto que se busca.

Imagen que contiene Texto

Descripción generada automáticamente

**Imagen 18. Excel resultante de la búsqueda (Microsoft, 2023).**

En conclusión, el desarrollo del RPA ha sido un logro significativo que ha culminado en una solución robusta y adaptable. Este sistema automatizado ofrece a la empresa la posibilidad de optimizar sus procesos, reducir errores y aumentar la eficiencia general. Con el RPA en funcionamiento, se espera que las tareas manuales y repetitivas sean reemplazadas por procesos automatizados, liberando tiempo y recursos para tareas de mayor valor agregado.

# **CONCLUSIONES**

El proyecto de implementación de un módulo RPA ha sido concluido con éxito, lo que brinda a la empresa Metaphorce, a través de Forte Innovation Consulting, la oportunidad de mejorar su proceso de búsqueda de perfiles. Esta implementación ha logrado reducir significativamente el tiempo dedicado a esta labor intensiva.

El proyecto ha cumplido con los objetivos clave, proporcionando un sistema integrado en la plataforma de Metaphorce para el personal de reclutamiento. Esta integración permite un acceso fluido y eficiente a las herramientas de búsqueda de perfiles, mejorando la experiencia y la eficiencia del equipo de reclutamiento.

El equipo de desarrollo se siente altamente satisfecho con los resultados obtenidos. Se ha enfocado en la aplicación de buenas prácticas de desarrollo de software, lo que ha resultado en un código de alta calidad, entregas puntuales y una buena documentación de los procesos. Estas prácticas han garantizado un producto final confiable y de calidad.

Por su parte, el equipo de reclutamiento ha encontrado esta herramienta extremadamente útil para facilitar sus tareas de búsqueda de perfiles. La herramienta ha mejorado su eficiencia y les ha permitido realizar las tareas de manera más efectiva.

En resumen, la implementación del módulo RPA ha sido un éxito para Metaphorce y Forte Innovation Consulting. Ha logrado optimizar el proceso de búsqueda y contratación de perfiles, reduciendo el tiempo empleado y mejorando la eficiencia del equipo de reclutamiento. El equipo de desarrollo se enorgullece de su trabajo y el equipo de reclutamiento ha encontrado en esta herramienta una solución valiosa para sus necesidades.

# **REFERENCIAS**

**Pressman, R.S.** (2010).*Ingeniería de software Un enfoque práctico.* (7ma ed.). Nueva York: Mc Graw Hill.

**Kaelble, S.** (2018). *Robotics Process Automation for Dummies (1ra ed.).* John Wiley & Sons, Limited, The Atrium, Southern Gate

**Rumbaugh et al.** (2007).*El lenguaje unificado de modelado manual de referencia. (2da ed.). Madrid: Person Education.*

**Duran. M.A.** (2023).*Aprendiendo git.* Leanpub.

**Haverbeke. M.**(2018). *Eloquent JavaScript. (3ra ed.).*

**Gamma et al.** (2003). *Patrones de Diseño Elementos de software orientado a objetos reutilizable.* Madrid: Person Education*.*

**Irrazábal E. Mascheroni m.**(2022). *Fundamentos de las pruebas continuas de software.*

**Listwon B. y Hanchett.** (2018). *Vue in action*. Estados Unidos de América: Manning Publications.

**Grinberg. M.** (2014). *Flask Web Development.* (1ra ed.). Estados Unidos de América: O’Reilly.

**Cone, M.** (2022). *The Markdown Guide.*