



Diseño e Implementación de un ERP con Automatización por Inteligencia Artificial y OCR para Aumentar la Independencia Tributaria de Pequeñas Empresas Chilenas.

Fernolds Alvarado Quispe

`contacto@efecompany.cl`

Resumen: En Chile, los procesos tributarios de pequeñas empresas presentan dependencia de contadores y un uso excesivo de tiempo en tareas como la preparación de declaraciones de IVA. Las soluciones disponibles en el mercado ofrecen funcionalidades relevantes; sin embargo, requieren conocimientos contables y tributarios para su uso.

El presente documento describe el diseño e implementación de una plataforma ERP que integra inteligencia artificial y reconocimiento óptico de caracteres (OCR) para automatizar la extracción y procesamiento de documentos contables y tributarios.

El objetivo principal fue evaluar en qué medida la solución desarrollada aumenta la automatización de procesos, mejora la precisión de registros y fomenta la independencia tributaria de los dueños de pequeñas empresas. La metodología de desarrollo utilizada fue Lean Startup, dado que existe evidencia sobre su efectividad para la construcción de productos mínimos viables y la validación iterativa de hipótesis. La validación se realizó mediante un diseño cuasi-experimental con dos grupos: contadores profesionales (Grupo A) que evaluaron la plataforma frente a métodos manuales y softwares comerciales, y dueños de empresa (Grupo B) que ejecutaron la declaración mensual de IVA de 25 empresas durante la primera semana de marzo de 2026 a través del agente Renata. El 100% de los cálculos fue completado sin errores y sin intervención contable en la etapa de cálculo, lo que evidenció una reducción efectiva de la dependencia del conocimiento contable especializado en el proceso tributario.

Palabras Clave: Inteligencia Artificial, Chile, Automatización de Procesos, Tributaria.



CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): ☐ Memoria o trabajo de título ☒ Tesis de Postgrado

Título del trabajo: Diseño e Implementación de un ERP con Automatización por Inteligencia Artificial y OCR para Aumentar la Independencia Tributaria de Pequeñas Empresas Chilenas

Nombre del candidato(a): Fernolds Bryan Alvarado Quispe

Carrera / Grado: Magíster en Tecnologías de la Información

Campus: San Joaquín **Departamento:** Informática

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, Jose Luis Marti Lara, en mi calidad de profesor(a) guía/director(a) del trabajo académico mencionado anteriormente **DEJO CONSTANCIA** que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución.

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL (marcar una opción)

☒ El trabajo **NO contiene** información que amerite confidencialidad y puede ser publicado de inmediato en repositorio con acceso abierto.

☐ El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad (**embargo**) por (**marcar una opción**):

☐ 6 meses ☐ 12 meses ☐ 2 años ☐ 3 años ☐ 5 años ☐ 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

4.- FIRMAS

Profesor(a) guía o director(a) de memoria o tesis:

Fecha: 18 - 08 - 2026

Firma:

Estudiante o Candidato(a):

Fecha: 17 de agosto de 2026 **Firma:**

Este formulario debe ser insertado como página 2 de la memoria o tesis, completado y firmado por estudiante y profesor(a) antes de la entrega en portal PRISMA de Biblioteca USM.

1 Introducción

1.1 Contexto

En Chile, en el año 2024, las pequeñas y microempresas fueron de 1.170.044 empresas, (ver Figura 1) [1]. Dentro de sus principales desafíos que enfrentan estas organizaciones es la gestión tributaria caracterizada por una alta complejidad normativa y una fuerte dependencia de profesionales tributarios para el cumplimiento de obligaciones fiscales [2].

Año Comercial	Tramo según ventas (5 tramos)	Número de empresas
2024	Sin Ventas/Sin Información	356.419
2024	Micro	909.469
2024	Pequeña	260.575
2024	Mediana	38.354
2024	Grande	17.988

Figura 1. Número de empresas. Fuente: SII 2024

La transformación digital en el ámbito contable – tributario ha impulsado el desarrollo de soluciones basadas en software, como Nubox [3], que buscan automatizar registros contables, conciliaciones y reportes financieros. Sin embargo, el uso de estas herramientas es eficaz para profesionales del área, pero no siempre resultan accesibles y aptos para dueños de negocio sin formación contable – tributaria. Adicionalmente los contadores requieren funciones de recopilar y ordenar gastos que no son registrados de manera automática en el SII (Servicio de Impuestos Internos) y para esto los contadores utilizan herramientas como por ejemplo Rindegastos [4], que permite a los usuarios a través de una foto ir registrando las compras realizadas.

1.2 Definición del problema

En Chile existen varios regímenes tributarios [5] de los cuales, en esta tesina, se enfocó en los regímenes Pro-Pyme General y Pro-Pyme Transparente, que tienen la posibilidad de llevar contabilidad simplificada, lo que significa que no es necesario que lleven todos los libros contables [5]. Dicha situación genera una oportunidad para automatizar tareas e informes que actualmente están semiautomatizados gracias al uso de sistemas ERP; sin embargo, como se indicó anteriormente aún se requieren expertos y conocimiento contable y tributario para cumplir la normativa vigente, lo cual se traduce en alta dependencia de contadores profesionales lo que limita la autonomía de los dueños de negocio. El problema se ve agravado porque gran parte de estas empresas aún recurre a procesos manuales, lo que produce ineficiencias, errores en los cálculos tributarios y pérdida de tiempo administrativo. Si bien no se logró encontrar información oficial sobre el número de rectificaciones de IVA, cabe mencionar que los errores al declararlo mal conllevan a multas [6], y reajustes que resulta ser un duro golpe para las pequeñas empresas que buscan hacer crecer su negocio.

1.3 Alternativas de solución existentes

La principal alternativa para la declaración de impuestos en Chile es la plataforma web del Servicio de Impuestos Internos, que centraliza documentación tributaria como facturas y boletas; sin embargo, depende de información enviada por terceros, lo que puede generar errores y sanciones. En paralelo, han surgido soluciones privadas orientadas a la automatización contable, como Nubox y Defontana ERP Pymes [7], que aún carecen de integración con IA y requieren intervención manual. Otras propuestas, como Databot [8], se enfocan en ser un agente de venta mediante WhatsApp, mientras que Rindegastos se limita a la gestión de gastos y viáticos.

A nivel internacional, herramientas como Veryfi [9] y QuickBooks [10] ofrecen funciones avanzadas de contabilidad y OCR, pero sin adaptación normativa ni lingüística al contexto chileno ni integración conversacional inteligente como la propuesta de trabajo.

1.4 Solución propuesta

La plataforma fue desarrollada bajo una arquitectura modular distribuida compuesta por tres sistemas interconectados. El módulo ERP Taxi gestionó la lógica contable, tributaria y laboral, exponiendo una API REST como núcleo central del sistema. El módulo Renata implementó un agente autónomo de inteligencia artificial capaz de procesar consultas en lenguaje natural y ejecutar operaciones contables mediante sub-agentes especializados, utilizando como canal de interacción una interfaz web propia con comunicación en tiempo real. El módulo Rasca automatizó la descarga del Registro de Compras y Ventas desde el portal del SII mediante automatización de navegador web.

Durante el desarrollo se implementaron y evaluaron dos componentes adicionales: un módulo RAG basado en circulares del SII y una integración conversacional vía WhatsApp, ambos descartados o pausados por razones técnicas y operacionales detalladas en la sección 3.6. El sistema utilizó agentes autónomos especializados en áreas tributarias, laborales y de soporte, coordinados por un orquestador central que determinó el flujo de respuesta según el contexto de cada conversación.

1.5 Objetivos

1.5.1 Objetivo general

El objetivo general de la investigación es diseñar e implementar una plataforma de asistencia contable y tributaria que integre un sistema de gestión empresarial con inteligencia artificial generativa y reconocimiento óptico de caracteres, evaluando en qué medida la automatización del cálculo tributario reduce la dependencia de los dueños de pequeñas empresas chilenas respecto del conocimiento contable especializado para la ejecución de procesos tributarios periódicos.

1.5.2 Objetivos específicos

- Implementar y evaluar un sistema de recuperación aumentada de información para automatizar la consulta de normativas tributarias vigentes, para determinar su viabilidad en el contexto del sistema desarrollado.
- Desarrollar una interfaz de interacción propia para facilitar el acceso a las funciones del agente autónomo desde un entorno cotidiano para los usuarios.
- Evaluar en qué medida el sistema reduce la participación del conocimiento contable especializado en la ejecución de la declaración mensual de IVA, midiendo la capacidad del dueño de empresa para completar el proceso con apoyo exclusivo del sistema.

1.6 Hipótesis y metodología de validación

La hipótesis es que **“La integración de un ERP con automatización basada en inteligencia artificial y reconocimiento óptico de caracteres reduce la dependencia de los dueños de pequeñas empresas chilenas respecto del conocimiento contable especializado para la ejecución de la declaración mensual de IVA”**.

En cuanto a la metodología de validación de la investigación es de tipo Experimental Aplicado por lo cual para la validación se realizó la “Construcción de Mínimo Producto Viable (MVP) y realización de pruebas de concepto según lo mencionado por F. J. Llamas [34].

2 Marco Teórico y Estado de Arte

2.1 Sistemas ERP y digitalización contable

Los sistemas de planificación de recursos empresariales (ERP) se han consolidado como herramientas fundamentales para integrar procesos contables, financieros y administrativos. En el ámbito de las Pymes los ERP permiten centralizar la gestión de facturación, inventarios y cumplimiento tributario. Estudios afirman que el uso de ERP en empresas chilenas es más favorable para estas, entregando mayor valor a comparación de tener contratado un equipo TI propio [16]. Sin embargo, estas mejoras y valor agregado requieren cierto grado de capacitación y conocimiento previo de este.

La primera alternativa para la declaración de impuestos y principal plataforma de uso tributario es la plataforma web del Servicio de Impuestos Internos. En esta se encuentra parte de la documentación tributaria de las empresas, como facturas, boletas y comprobantes de pago. Sin embargo, la plataforma se alimenta con información enviada por proveedores de medios de pagos y el uso de la propia plataforma, lo cual abre la posibilidad a que en ciertas ocasiones los proveedores envían información **incorrecta**, y esto se traduce en propuestas erróneas lo que desembocan en multas e intereses.

Tal como se menciona en la sección 1.3 de este informe, existen varios softwares y ERPs en el mercado chileno que se enfocan en la parte financiera y que para su uso requiere conocimiento contable y tributario. En el ámbito internacional se han desarrollado plataformas de integración de Whatsapp con OCR como ser Verify [9]. Sin embargo, no tiene mayores funcionalidades contables y tributarias que la de extraer información de la imagen al igual que lo hace Rindegastos. Otro de los grandes referentes anglosajón es Quickbooks [10], que cuenta con bastantes funciones de contabilidad, pero de igual manera para el uso se requiere conocimientos contables y tributarios, tampoco se evidencia uso de IA de la misma manera que lo propuesto en este estudio. No obstante, la mayoría de estas soluciones se diseñan para mercados anglosajones y carecen de adaptación normativa y lingüística al contexto chileno.

2.2 Inteligencia Artificial

Los inicios en 1943, conocidos como la "Primera Primavera de la IA", constituyen el período fundacional de la IA como disciplina científica, marcada por una serie de hitos teóricos y experimentales que sentaron las bases para su desarrollo.

Un punto de inflexión clave fue la propuesta del modelo de neurona artificial de McCulloch-Pitts en 1943 [17], el cual, al formalizar matemáticamente la función de una neurona biológica, proporcionó el andamiaje para el diseño de sistemas de procesamiento de información inspirados en el cerebro.

La consolidación del campo se produjo con el taller de Dartmouth en 1956 [18], un evento que no solo acuñó el término IA sino que también estableció sus objetivos como la simulación del pensamiento humano, catalizando un fervor por la investigación que se materializó en innovaciones como el perceptrón de Rosenblatt en 1958 [19], este último un algoritmo de aprendizaje supervisado que demostró la viabilidad de entrenar sistemas para reconocer patrones, aunque su éxito inicial se vio limitado a problemas linealmente separables.

Estos avances conceptuales, combinados con las ideas teóricas de la máquina de Turing de 1950 [20] sobre la computabilidad, definieron el marco de una nueva ciencia, posicionando la IA no solo como una rama de la informática, sino como un campo interdisciplinario con profundas implicaciones para la filosofía, la lógica y la cognición.

Y los académicos más actuales como J. McCarthy la IA [21] indican que esta “es la ciencia e ingeniería de hacer máquinas inteligentes, especialmente programas de computadora inteligentes. Está relacionada con la tarea similar de usar computadoras para entender la inteligencia humana, pero la IA no tiene que limitarse a métodos que sean biológicamente observables”.

2.2.1 Aprendizaje Automático

Una de las definiciones más citadas en el campo del aprendizaje automático es la propuesta por Tom M. Mitchell en su libro *Machine Learning* (1997) [23], donde establece una formalización concisa del concepto de aprendizaje. Según Mitchell, se considera que un programa aprende si, al enfrentarse a una tarea (T) y con una determinada medida de rendimiento (P), su desempeño en dicha tarea mejora con la experiencia (E). Esta definición trasciende la simple programación de reglas y subraya que el aprendizaje es un proceso de mejora iterativa y empírica, donde un sistema perfecciona su capacidad para resolver un problema a medida que procesa más datos. La relevancia de esta formulación reside en que proporciona un marco claro para la evaluación de los algoritmos de aprendizaje, permitiendo que el progreso en la disciplina sea medible y comparable, sentando las bases para el desarrollo de los sistemas modernos de IA.

2.2.2 Aprendizaje Profundo

Como se visualiza en la Figura 2, entre la comparativa de inteligencia artificial, aprendizaje automático y aprendizaje profundo, se puede indicar que aprendizaje profundo [22] es solo una parte de todo lo que implica aprendizaje automático. Este documento se enfoca más en este punto que en las otras funcionalidades y tareas de aprendizaje automático ya que el presente estudio se utilizan herramientas de OCR y NLP.

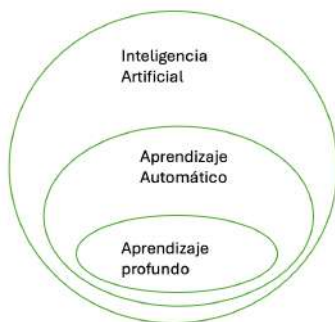


Figura 2. Relación entre inteligencia artificial, aprendizaje automático y aprendizaje profundo

El aprendizaje profundo se define como una rama del aprendizaje automático que utiliza modelos computacionales compuestos por múltiples capas de procesamiento para aprender representaciones de datos con múltiples niveles de abstracción [24]. Estos métodos han mejorado drásticamente el estado del arte en campos como el reconocimiento de voz, de objetos visuales. A diferencia de los métodos de aprendizaje automático convencionales, que requieren una cuidadosa ingeniería de características por parte de expertos humanos para transformar los datos brutos, las redes de aprendizaje profundo aprenden automáticamente estas representaciones a partir de los datos utilizando un procedimiento de aprendizaje de propósito general, como la retropropagación. Esto ha permitido que el aprendizaje profundo resuelva problemas que habían resistido a la comunidad de IA durante muchos años, como el reconocimiento de imágenes, de voz y el procesamiento del lenguaje natural, al requerir mucho menos trabajo manual.

Un área clave de aplicación del aprendizaje profundo es el OCR y el procesamiento del lenguaje natural (NLP). Las redes neuronales convolucionales (ConvNets) han logrado un éxito notable en la lectura de documentos y el reconocimiento de caracteres escritos a mano. Para tareas secuenciales como el lenguaje, las redes neuronales recurrentes (RNNs) son más adecuadas, ya que procesan la entrada un elemento a la vez, manteniendo un "vector de estado" que contiene información sobre la historia de la secuencia. Esta capacidad permite a las RNNs, a través de modelos de lenguaje probabilísticos, aprender representaciones vectoriales de palabras que codifican características semánticas. OCR Y NLP son tecnologías de gran valor para la automatización de tareas

y funciones diarias contables tributarias, ya que un contador habitual utiliza la vista para identificar registros y clasificar cuentas contables que permiten obtener resultados tributarios.

2.2.3 IA generativa

Con el nacimiento de Open AI [25] es cuando se popularizó el término “IA Generativa”. Según J. Casar “es el conjunto de métodos y aplicaciones capaces de generar contenidos con características indistinguibles de las que produciría un ser humano” [26]. Así, es una gran herramienta para el desarrollo de este documento, ya que su uso entrega herramientas necesarias para la automatización e independencia de las actividades y funciones tributarias chilenas.

Posterior al surgimiento de Open AI nacieron otros competidores, como Claude, Gemini, Llama, Grok y DeepSeek, los que permiten generar textos similares a los que un ser humano puede crear al momento de comunicarse con otro ser humano. Su gran beneficio, en el contexto actual, es ser aplicado mediante agentes autónomos o *bots* que permiten una comunicación más fluida, a diferencia de los *bots* programados que solo permiten comunicarse a través de ciertas opciones predefinidas, como por ejemplo uno de soporte al cliente que ofrece unas cuantas opciones al usuario, limitando de esta manera una conversación más fluida y natural con la máquina.

Un problema muy habitual de las IAs generativas son las alucinaciones. Estas, inicialmente, en las primeras versiones eran más recurrentes; sin embargo, debido a los avances de nuevos modelos y estudios académicos al respecto, su porcentaje de ocurrencia fue disminuyendo, entregando de esta manera resultados más realistas y cercanos a lo que un humano puede generar.

2.2.4 Modelos de Lenguaje Grande

Un modelo de lenguaje grande (LLM) es un modelo computacional de aprendizaje profundo, basado predominantemente en la arquitectura *Transformer*, que ha sido entrenado con datos textuales a escala masiva. Estos modelos funcionan aprendiendo patrones estadísticos, sintácticos y semánticos del lenguaje, lo que les permite generar, comprender, traducir y resumir texto con un alto grado de coherencia y fluidez. Su gran tamaño hace que requiera técnicas de compresión como para reducir su número de parámetros y así disminuir los recursos necesarios para su uso en tiempo real [27].

Los LLMs, han demostrado una capacidad sin precedentes para comprender y generar texto en lenguaje natural. Aplicados al dominio contable y tributario, permiten resolver consultas normativas, explicar procedimientos y asistir en la clasificación de documentos [28]. Esto se traduce en que el uso e implementación de IA en contabilidad tributaria tiene un gran potencial de entrega de valor y mejora de procesos y generación de reportes; no obstante, su aplicación al contexto contable chileno es aún incipiente lo que justifica la pertinencia de este estudio.

2.3 Recuperación Aumentada de Información (RAG)

Los LLMs, a pesar de sus vastos parámetros que almacenan conocimiento, enfrentan desafíos significativos como lo ya mencionado y la falta de experiencia en dominios específicos. Para mitigar estos problemas ha surgido RAG, que complementa a los LLMs con una base de datos de conocimiento externa.

Un sistema RAG [29] consta de tres componentes principales: un *retriever* que busca información relevante, un *generator* (LLM) que produce la respuesta y las *fusions* que combinan la información recuperada con la entrada del usuario (ver Figura 3). Las técnicas de fusión pueden ser de tres tipos: por consulta, que concatena texto o características con la entrada original; por *logits*, que ajusta las probabilidades de salida; y latente, que integra la información en las capas ocultas del modelo.

En el ámbito de la **contabilidad tributaria**, el RAG es fundamental para abordar los desafíos de la naturaleza altamente específica y en constante cambio de la legislación fiscal. El sistema RAG debe ser entrenado con una base de datos de conocimiento que incluya normativas fiscales, precedentes legales, circulares del Servicio de Impuestos Internos y manuales de procedimiento. Al recibir una consulta sobre un caso tributario complejo, el *retriever* busca y recupera la información legal más actualizada y relevante; posteriormente el *generator* (LLM) utilizando las técnicas de fusión para integrar esta información, genera respuestas precisas evitando la alucinación de datos o la desactualización.

Para lograr una evaluación adecuada se evaluaron con las siguientes fórmulas:

$$\text{Precisión} = \frac{TP}{TP + FP}$$

Formula 1. Precisión de respuestas correctas

$$\text{Recall} = \frac{TP}{TP + FN}$$

Formula 2. Recall

$$F1 = \frac{2 \times (\text{Precisión} \times \text{Recall})}{(\text{Precisión} + \text{Recall})}$$

Formula 3. F1 – score

donde:

- *TP* (True Positives): número de respuestas del modelo que fueron correctas respecto al estándar gold.
- *FP* (False Positives): número de respuestas incorrectas dadas por el modelo.

Estas fórmulas ayudan a evaluar sistemas RAG en relación a un modelo LLM que no utiliza RAG, vale decir un *baseline*.

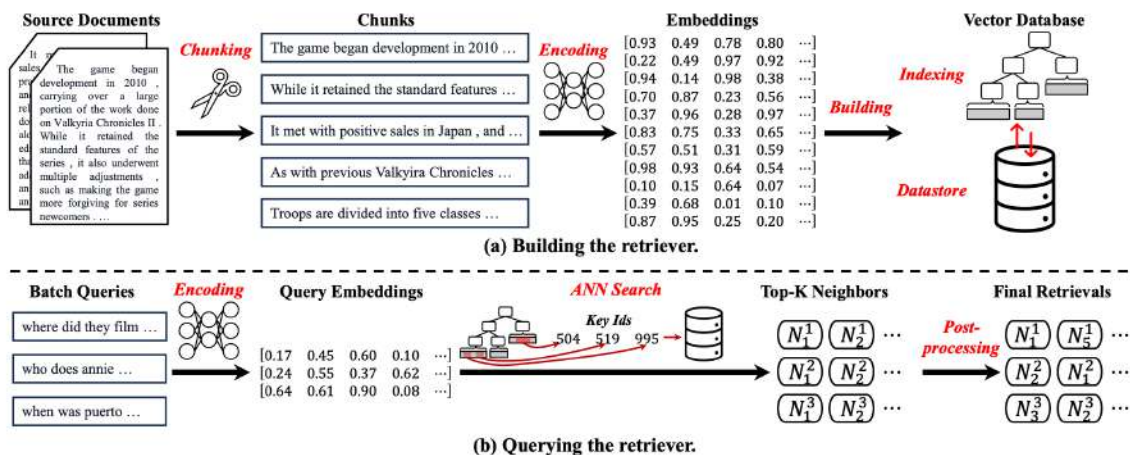


Figura 3. Modelo RAG [30]

Métricas de similitud para RAG

Los sistemas de RAG se basan en la búsqueda de los "vecinos más cercanos" para encontrar información relevante. Y la eficacia de esta búsqueda depende de las métricas de distancia o similitud utilizadas. La elección de una métrica impacta directamente en la calidad de los resultados, ya que determina cuán relevante es un fragmento de texto para una consulta. Existen tres métricas principales, según S. Wu et al. [30]: la similitud del coseno, la distancia euclidiana y la distancia de Manhattan, las cuales se definen a continuación.

- Similitud del Coseno: Mide el ángulo entre dos vectores para determinar su semejanza. A diferencia de otras métricas de distancia, la similitud del coseno no se ve afectada por la magnitud de los vectores, lo que es útil en el NLP donde los vectores de *embeddings* representan el significado semántico de un texto y la longitud del texto no debe afectar la relevancia. Un valor de similitud del coseno cercano a 1 indica que los vectores tienen una dirección muy similar lo que se traduce en una alta relevancia semántica.
- Distancia Euclidiana: Mide la distancia en línea recta entre dos puntos en un espacio multidimensional. Es una métrica de distancia simple, aunque es intuitiva para espacios de baja dimensionalidad; su efectividad puede disminuir significativamente en el contexto de los *embeddings* de lenguaje que, a menudo, operan en espacios de muy alta dimensionalidad.
- Distancia de Manhattan: Mide la distancia entre dos puntos sumando las diferencias absolutas de sus coordenadas; esto significa que las diferencias de la distancia manhattan no miden una línea recta, sino el camino recorrido a lo largo de los ejes.

2.4 Agente Autónomo

El surgimiento de los agentes autónomos basados en LLM representa un avance significativo en la consecución de la IA general [31]. A diferencia de los agentes tradicionales que operaban con conocimientos limitados en entornos aislados, los agentes con LLM aprovechan el vasto conocimiento adquirido de Internet para exhibir capacidades de toma de decisiones similares a las humanas.

Para que el agente autónomo logre cumplir con sus objetivos iniciales de entregar información adecuada a los usuarios, se pueden utilizar distintas técnicas para su adecuada capacitación, algunas de las cuales se presentan a continuación.

- Adquisición de Capacidades con Ajuste Fino: Esta estrategia implica el uso de conjuntos de datos específicos para una tarea con el fin de ajustar los parámetros del LLM. Los conjuntos de datos pueden ser creados a partir de anotaciones humanas, por otros LLMs o a partir de datos del mundo real.
- Adquisición de Capacidades sin Ajuste Fino: Estos métodos no modifican los parámetros del LLM, sino que se basan en técnicas de *prompt engineering* o ingeniería de mecanismos. Son útiles tanto para LLMs de código abierto como para los de código cerrado.

2.5 OCR en procesos contables

OCR es una tecnología clave para la digitalización de boletas facturas y documentos tributarios. Su precisión ha mejorado significativamente con la aplicación de modelos de aprendizaje profundo. En Chile, donde la boleta electrónica se masificó en 2021, el OCR permite integrar información tributaria en sistemas ERP sin necesidad de digitación manual [20].

Es un sistema compuesto por dos etapas principales: la detección de texto y el reconocimiento de texto. Las metodologías iniciales trataban ambas tareas de forma independiente, pero las investigaciones recientes han tendido a integrar los modelos de detección y reconocimiento en un marco unificado para mejorar la eficiencia y precisión del proceso. En el contexto de la detección de texto en imágenes, el OCR no solo es fundamental para identificar las palabras, sino también para que el modelo de lenguaje pueda comprender su significado y su relación con los objetos visuales que incluyen texto de la escena.

2.6 Arquitectura de software

La arquitectura de software define la estructura fundamental de un sistema, organiza sus componentes y determina cómo se comunican entre sí. Según la literatura académica, existen diversos estilos y patrones que ofrecen soluciones probadas a problemas recurrentes de diseño. Entre los modelos más utilizados se encuentran la arquitectura en capas, la arquitectura modular y la arquitectura de microservicios. El estilo en capas organiza la aplicación en niveles horizontales, como la presentación, la lógica de negocio y la persistencia de datos, lo que facilita la separación de responsabilidades y el mantenimiento del sistema. Por su parte, la arquitectura modular se fundamenta en la división del sistema en unidades funcionales independientes y cohesivas, lo que permite el desacoplamiento de procesos y la distribución de componentes en distintos entornos de ejecución. Finalmente, la arquitectura de microservicios divide la aplicación en servicios pequeños e independientes que se comunican a través de interfaces ligeras, lo que potencia la escalabilidad, la resiliencia y la flexibilidad tecnológica del conjunto [33].

2.7 Metodología de trabajo

La metodología *Lean Startup* es un enfoque de emprendimiento que se centra en la eficiencia y el aprendizaje validado. Propone que, en lugar de crear una solución directamente, se lance una *Startup* como un experimento para validar un modelo de negocio antes de invertir grandes cantidades de tiempo, recursos y esfuerzo [34]. El objetivo principal de la metodología es descubrir rápidamente qué es lo que los usuarios realmente necesitan, minimizando la incertidumbre y el riesgo de fracaso. Se basa en un ciclo iterativo de tres pasos: crear un producto, medir los resultados y aprender de ellos; a continuación, se detallan las etapas de la metodología.

- **Crear:** Al inicio del proyecto, no se tiene información completa para desarrollar un producto perfecto que se ajuste a las necesidades de los usuarios. Por eso, la metodología *Lean Startup* propone crear un PMV. El PMV es una versión del producto con las funcionalidades esenciales para recopilar la mayor cantidad de aprendizaje validado sobre los clientes con el mínimo esfuerzo y en el menor tiempo posible. Su objetivo no es impresionar al usuario en general, sino probar las hipótesis fundamentales del negocio con un grupo pequeño de *early adopters* o clientes pioneros que están dispuestos a adoptar el producto en sus fases iniciales.
- **Medir:** Una vez que el PMV está en manos de los primeros usuarios, el siguiente paso crucial es medir su respuesta y comportamiento. El principal desafío de esta fase es determinar cómo reaccionan los usuarios para obtener datos fiables que permitan tomar decisiones informadas.
- **Aprender:** El objetivo final de todo el circuito *Lean Startup* es generar aprendizaje validado, dado que la información recopilada en la fase de medición ayuda a decidir si la propuesta de solución es viable. Si los datos indican que las hipótesis iniciales son correctas, el equipo puede perseverar y seguir desarrollando versiones incrementales del producto. Si, por el contrario, las hipótesis no se validan es necesario pivotar que implica realizar un cambio significativo en uno o varios componentes del modelo para probar una nueva hipótesis. El pivotar no es un fracaso, sino una corrección estructurada.

2.8 Metodología de desarrollo

Existen distintas metodologías de desarrollo de software. Una de ellas sigue un ciclo de vida evolutivo, que corresponde a un enfoque de desarrollo de software que, en lugar de seguir un proceso lineal y rígido, se centra en la iteración y la mejora continua. Se caracteriza por la entrega de versiones funcionales y parciales del software en ciclos cortos. Cada ciclo se basa en la retroalimentación del cliente y de los usuarios para evolucionar el producto. En la metodología de ciclo de vida evolutivo [35], el desarrollo no se detiene después de la primera entrega; el software se adapta y crece con el tiempo integrando nuevas funcionalidades y mejoras. Este método es útil para proyectos donde los requisitos iniciales no están completamente claros o pueden cambiar, ya que se pueden hacer ajustes a medida que avanza el proyecto. Siguiendo con la metodología, cada módulo fue sometido a pruebas controladas y los errores detectados durante estos ciclos fueron corregidos, y

en paralelo se incorporaron mejoras sugeridas tanto por pruebas técnicas como por retroalimentación de usuarios (contadores y dueños de pequeñas empresas).

2.9 Estado del arte en soluciones contables asistidas por IA

A nivel internacional ya existen varios estudios relacionados a este tema. Por ejemplo, [36] evidencia que la IA otorga mayor automatización en la entrada de registros contables, de igual manera se indica la posibilidad de utilizar OCR para la digitalización de documentos. Sin embargo, también se resalta que la implementación de IA en empresas pequeñas y medianas representa un desafío mayor, ya que estas tecnologías generalmente son utilizadas por grandes empresas por temas de costos y necesidad de conocimientos previos para el uso e implementación en los procesos financieros y tributarios. No obstante, la mayoría de estas soluciones están diseñadas para mercados anglosajones y carecen de adaptación normativa chilena y no integran canales de mensajería instantánea. En Chile, a modo de ejemplo, el trabajo de J. Sepúlveda Rivera [37], busca probar hasta qué grado se pueden automatizar ciertas tareas contables, y cuáles son sus limitaciones y alcances de funciones automatizables a fecha de hoy. Estudios similares en Chile pueden evidenciar en el 2024 por Matías Cartes [38], donde indica que para entonces no se puede utilizar la IA para tareas complejas en la auditoría y lo que sí se puede hacer es utilizarla para potenciar y acelerar el trabajo. Si bien no está relacionado directamente con el trabajo de este documento, para el autor pareció ser relevante mencionar que hay ciertos trabajos que buscan analizar la implementación de IA en distintas áreas y proceso de las ramas de la contabilidad y similares.

2.10 SUS

System Usability Scale [11] es un cuestionario estandarizado de 10 ítems desarrollado por Brooke en 1986, ampliamente utilizado para la medición de la usabilidad percibida de sistemas interactivos. Se basa en una escala tipo Likert de cinco puntos que permite obtener un puntaje global de 0 a 100 donde valores superiores a 70 se interpretan como aceptables o buenos niveles de usabilidad. Su principal fortaleza radica en que constituye una herramienta validada y de bajo costo de aplicación y con alta confiabilidad incluso en muestras pequeñas lo que lo hace especialmente adecuado para estudios de prototipos o pruebas de concepto.

2.11 NASA-TLX

NASA-TLX es un cuestionario ampliamente utilizado en estudios de ergonomía y factores humanos para evaluar la carga de trabajo subjetiva percibida por los usuarios durante la ejecución de tareas. Fue desarrollado por el NASA Ames Research Center y se compone de seis dimensiones: demanda mental, demanda física, demanda temporal, esfuerzo, rendimiento percibido y nivel de frustración. Cada dimensión se mide en una escala de 0 a 100, y en conjunto proporciona un índice global de carga de trabajo que permite identificar el nivel de exigencia cognitiva y emocional asociado al uso de un sistema [14].

3 Propuesta de Solución

En este capítulo se describen los aspectos más relevantes del diseño e implementación de la plataforma propuesta, la cual estuvo compuesta por tres sistemas interconectados: un ERP contable-tributario-laboral denominado Taxi, un agente autónomo de inteligencia artificial denominado Renata, y un servicio de automatización de consultas tributarias al SII denominado Rasca. No se abordaron en profundidad aspectos de código que no aportaran valor al lector.

3.1 Arquitectura General del Sistema

La plataforma adoptó una arquitectura modular distribuida en tres componentes independientes que operaron de forma coordinada, como se aprecia en la Figura 4. Esta decisión respondió a la necesidad de separar

responsabilidades, facilitar el mantenimiento individual de cada servicio y permitir la escalabilidad de cada componente según su demanda particular.

- **Taxi:** ERP *backend* que expuso una API REST como núcleo de la lógica contable, tributaria y laboral. Fue el orquestador central de datos y procesamiento.
- **Renata:** Agente autónomo de inteligencia artificial que interactuó con los usuarios finales mediante una interfaz web propia y, en fase experimental, mediante WhatsApp. Consumió la API de Taxi para ejecutar operaciones contables.
- **Rasca:** Servicio especializado en la extracción automatizada de documentos tributarios desde el portal del Servicio de Impuestos Internos mediante automatización de navegador web. Fue consumido directamente por Taxi para la descarga del Registro de Compras y Ventas (RCV).

La comunicación entre los tres sistemas se realizó mediante interfaces REST autenticadas. Taxi actuó como intermediario entre Renata y Rasca, de modo que Renata nunca consumió directamente a Rasca, sino que lo hizo a través de los *endpoints* de Taxi.

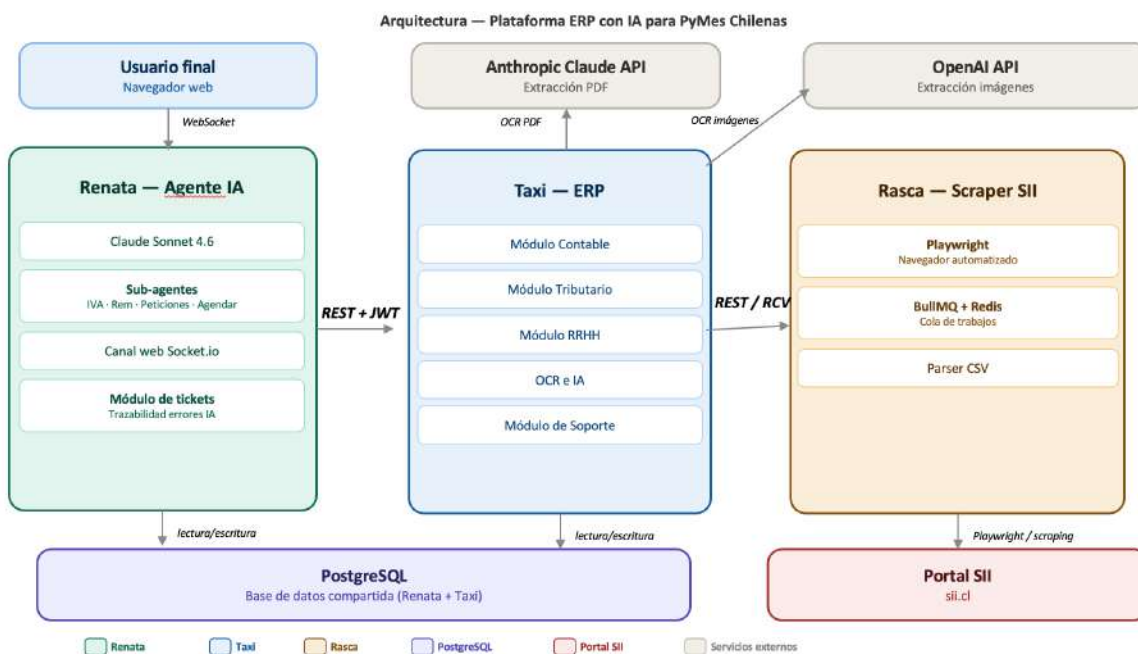


Figura 4 Arquitectura de solución propuesta

3.2 Módulo ERP Contable-Tributario-Laboral (Taxi)

Taxi está desarrollado en Javascript, Node.js con Express, que entrega una interfaz REST para orquestar las operaciones de lógica de negocio. La persistencia de datos se implementó mediante PostgreSQL a través de la biblioteca *pg-promise*, seleccionada por su robusta gestión de transacciones y optimización del rendimiento mediante agrupación de conexiones.

La autenticación y autorización de solicitudes se resolvieron mediante *JSON Web Tokens (JWT)*, lo que garantizó que toda interacción con la API fuera debidamente verificada. Adicionalmente, se implementó un esquema de autenticación de servicio a servicio, mediante el cual Renata obtuvo un *token* especial para operar como cliente autorizado de Taxi.

Las principales funcionalidades del sistema incluyeron:

- **Módulo Contable:** Registro de asientos con múltiples cuentas, carga masiva de compras y ventas desde archivos CSV o Excel, generación de libro diario, libro mayor y balance de ocho columnas, y cierre mensual y anual.
- **Módulo Tributario:** Generación de propuesta de declaración de IVA (Formulario 29), procesamiento de declaraciones juradas DJ1887, DJ1879, DJ1948 y DJ1835, y generación del Formulario 22 de declaración de renta.
- **Módulo RRHH y Remuneraciones:** Gestión del ciclo completo de remuneraciones incluyendo liquidaciones de sueldo, contratos de trabajo, finiquitos, licencias médicas, retención de honorarios, integración con Previred para cotizaciones previsionales e indicadores previsionales mensuales. Los cálculos laborales se implementaron conforme a la normativa chilena vigente, incluyendo gratificación, indemnización por años de servicio y tarifas diarias y horarias.
- **Módulo de Procesamiento de Documentos:** Extracción de datos mediante OCR e IA desde boletas, facturas, certificados Previred, contratos de arriendo y formularios tributarios en formato PDF e imagen.
- **Módulo de Soporte (Taxi):** Sistema de *tickets* para gestionar solicitudes especiales que la plataforma no logró automatizar, tales como trámites presenciales ante el SII, municipalidades u otras instituciones. Este módulo permitió a los contadores recibir, gestionar y resolver dichas solicitudes desde la misma plataforma.

3.3 Automatización de Consultas al SII — Rasca

Rasca es un servicio especializado desarrollado en TypeScript con Fastify, cuya función principal es automatizar la descarga y procesamiento del Registro de Compras y Ventas (RCV) desde el portal web del Servicio de Impuestos Internos, utilizando automatización de navegador web mediante Playwright. El proceso de extracción comprendió las siguientes etapas: autenticación en el portal SII con las credenciales del contribuyente, navegación automatizada hacia el módulo RCV, selección del período tributario, detección y descarga del archivo CSV generado por el SII, y posterior *parseo* y almacenamiento en PostgreSQL. Dado que las operaciones de *scraping* podían extenderse en el tiempo, se implementó un sistema de procesamiento asíncrono mediante BullMQ sobre Redis, que permitió encolar trabajos de descarga y consultar su estado de avance mediante un identificador único. Las credenciales de los contribuyentes se almacenaron cifradas para garantizar su seguridad.

3.4 Funcionalidad OCR e Inteligencia Artificial

La extracción automatizada de información desde documentos se implementó mediante dos proveedores de inteligencia artificial, seleccionados según el tipo de documento a procesar. Para documentos en formato PDF, se utilizó el modelo *Claude Sonnet* de Anthropic [42], el cual permitió el procesamiento nativo de archivos PDF preservando la estructura tabular de los documentos. Esta capacidad resultó especialmente valiosa para certificados Previred, formularios tributarios y contratos, cuya información se organizó en tablas complejas. Los extractores implementados incluyeron: extracción de datos de certificados Previred, contratos de arriendo, información de colaboradores, Formulario F-22 y datos de socios. Para documentos en formato imagen, se utilizó el modelo *GPT-4o-mini* de OpenAI mediante codificación *base64*, principalmente para el procesamiento de boletas físicas digitalizadas.

En ambos casos, el flujo de procesamiento siguió el mismo patrón: el *frontend* envió el documento al *backend* de Taxi, el controlador correspondiente invocó al extractor de IA con un *prompt* especializado, la respuesta fue



limpiada y estructurada en formato JSON, y los datos fueron insertados en PostgreSQL. Esta estandarización permitió una validación consistente de campos como RUTs, fechas y montos.

3.5 Agente Autónomo de Inteligencia Artificial — Renata

Renata es un agente autónomo desarrollado en Node.js que utilizó el modelo *Claude Sonnet 4.6* de Anthropic mediante el SDK oficial, implementando el protocolo *Model Context Protocol* (MCP) para la invocación estructurada de herramientas externas.

3.5.1 Arquitectura del Agente

El agente operó mediante un orquestador central (*agentOrchestrator*) que coordinó el flujo de procesamiento de cada mensaje recibido. Dicho flujo incluyó: recepción y normalización del mensaje según el canal de origen, verificación del usuario en la base de datos compartida con Taxi, procesamiento de audio mediante ElevenLabs en caso de mensajes de voz, ejecución del agente Claude con historial de conversación, enrutamiento hacia el sub-agente especializado correspondiente, formateo de la respuesta y envío al usuario.

3.5.2 Sub-agentes Especializados

El sistema implementó cuatro sub-agentes especializados, cada uno con acceso a herramientas MCP específicas que invocaban *endpoints* de la API de Taxi:

- ***BalanceIvaAgent***: generación de balance de IVA, cálculo del monto a pagar, generación de enlaces de pago y verificación del estado de declaración.
- ***RemuneracionesAgent***: procesamiento de liquidaciones de sueldo, gestión de contratos y cálculo de finiquitos.
- ***OtrasPeticonesAgent***: gestión de solicitudes contables especiales, creación de *tickets* de soporte y consultas de cumplimiento tributario.
- ***AgendarAgent***: programación de reuniones con contadores mediante integración con Google Calendar.

3.5.3 Canal de Interacción Web

El canal de producción de Renata fue una interfaz web propia desarrollada con React y comunicación en tiempo real mediante Socket.io. Este canal permitió mantener trazabilidad completa de las conversaciones, lo que resultó fundamental para la detección de errores del agente y la mejora continua del sistema.

3.5.4 Módulo de Soporte de Renata

Se implementó un módulo de soporte de *tickets* específico para Renata, cuyo propósito fue permitir a los usuarios marcar respuestas incorrectas entregadas por el agente autónomo. Este mecanismo de *feedback* permitió al equipo de desarrollo identificar patrones de error, ajustar los *prompts* del sistema y mejorar iterativamente la precisión de las respuestas de la IA.

3.6 Decisiones de Diseño y Pivotes Metodológicos

Durante el proceso de desarrollo, se implementaron y evaluaron dos componentes que fueron posteriormente descartados o pausados, en aplicación del ciclo *Construir* → *Medir* → *Aprender* de la metodología *Lean Startup*.

Descarte del módulo RAG: Se implementó un sistema de recuperación aumentada de información basado en 25 circulares del SII, con *embeddings* generados mediante *text-embedding-ada-002* y almacenamiento vectorial en PostgreSQL. Las pruebas evidenciaron limitaciones significativas ante consultas fuera del dominio indexado, generando respuestas incorrectas que afectaron la confianza en el sistema. Adicionalmente, mantener actualizada una base normativa completa del SII y la Dirección del Trabajo representó un costo operacional elevado e inviable. Se decidió reemplazarlo por extracción directa mediante LLM, cuya capacidad de razonamiento general resultó más robusta para el dominio tributario chileno.

Pausa del canal WhatsApp: Se implementó una integración conversacional mediante la biblioteca de código abierto *whatsapp-web.js*. Las pruebas demostraron que el agente autónomo operó correctamente en términos funcionales; sin embargo, se identificaron dos problemas críticos: en primer lugar, Meta interrumpió deliberadamente la compatibilidad de bibliotecas no oficiales para promover el uso de su API de pago, lo que generó una carga de mantenimiento continua con actualizaciones de parches; en segundo lugar, el canal WhatsApp no permitió implementar un mecanismo de trazabilidad de errores del agente, impidiendo la detección y corrección sistemática de respuestas incorrectas. Se decidió pausar esta integración y priorizar el canal web propio como entorno de producción, donde la trazabilidad completa permitió alimentar el proceso de mejora continua del agente.

3.7 Interfaz de usuario Renata

La interfaz de usuario de Renata corresponde a una aplicación web de conversación en tiempo real, accesible desde cualquier navegador. El diseño adoptó un estilo de mensajería instantánea con fondo oscuro, donde el usuario interactuó con el agente mediante lenguaje natural, sin necesidad de conocimientos contables ni tributarios previos.

3.7.1 Lógica de Funcionamiento

Renata opera mediante un flujo de procesamiento secuencial coordinado por un orquestador central, cuyas etapas se describen a continuación:

- **Recepción y normalización del mensaje:** Cada mensaje es recibido desde el canal web es normalizado por un adaptador de mensajes (*messageAdapter*), el cual estandariza su formato independientemente del tipo de contenido (texto o audio).
- **Verificación del usuario:** El sistema consulta la base de datos compartida con Taxi para verificar la identidad del remitente y determinar su rol, distinguiendo entre clientes registrados y prospectos.
- **Procesamiento de audio:** En caso de que el mensaje correspondiera a un archivo de voz, el módulo de audio procesó la entrada mediante ElevenLabs para convertirla a texto antes de continuar con el flujo.
- **Ejecución del agente Claude:** El mensaje, junto con el historial completo de la conversación, fue enviado al modelo *Claude Sonnet 4.6* de Anthropic. El modelo analizó la intención del usuario y determinó cuál de los cuatro sub-agentes especializados debía atender la solicitud: *BalanceIvaAgent*, *RemuneracionesAgent*, *OtrasPeticonesAgent* o *AgendarAgent*.
- **Invocación de herramientas MCP:** El sub-agente seleccionado ejecutó las herramientas necesarias mediante el protocolo MCP, las cuales realizaron llamadas a los *endpoints* correspondientes de la API de Taxi para obtener datos contables, generar documentos o registrar operaciones.
- **Formateo y entrega de la respuesta:** La respuesta generada por Claude fue procesada por un módulo de formateo que eliminó caracteres no compatibles con el canal web, como *emojis* o sintaxis *markdown* no soportada, garantizando una presentación clara y consistente. En caso de requerirse respuesta de voz, el texto fue convertido nuevamente a audio mediante ElevenLabs.

3.7.2 Flujo de Pago y Cálculo de IVA

Uno de los flujos más relevantes implementados correspondió al cálculo y pago del IVA mensual, el cual ilustró de forma clara la integración entre Renata y Taxi. El flujo se desarrolló de la siguiente manera.

- **Paso 1 — Solicitud en lenguaje natural:** El usuario redactó su solicitud en lenguaje coloquial, sin necesidad de formularios ni menús. En el caso documentado, el mensaje fue: *"hola Renata necesito pagar el IVA de noviembre de 2025"*.
- **Paso 2 — Procesamiento por el agente:** Renata identificó la intención del usuario y enrutó la solicitud hacia el sub-agente *BalanceIvaAgent*, el cual invocó mediante herramientas MCP los *endpoints* correspondientes de la API de Taxi. Estos calcularon el balance de IVA del período solicitado y generaron un enlace de pago único asociado a la empresa del usuario.
- **Paso 3 — Respuesta del agente:** Renata entregó al usuario el monto adeudado (\$81.830 para el período de noviembre de 2025) junto con el enlace de pago generado dinámicamente por Taxi, sin que el usuario debiera acceder al sistema ERP directamente.
- **Paso 4 — Página de pago:** Al acceder al enlace entregado por Renata, el usuario fue dirigido a una página de pago generada por Taxi. Dicha página presentó de forma clara el período tributario, la razón social y RUT de la empresa, el monto a pagar y los datos bancarios para realizar la transferencia. Adicionalmente, la página ofreció la opción de cargar el comprobante de pago una vez realizada la transacción.

3.8 Interfaz de usuario Taxi

La interfaz de Taxi corresponde a una aplicación web desarrollada en React, orientada exclusivamente al uso del contador tributario. A diferencia de Renata, cuyo canal de interacción es conversacional y dirigido al dueño de la empresa, Taxi presenta un *panel* de administración estructurado por módulos funcionales, que permite al contador ejecutar operaciones contables, tributarias y laborales de forma directa y eficiente. El acceso a la plataforma es autenticado mediante credenciales únicas por empresa, mostrando en el encabezado la razón social de la organización activa, lo que permite al contador operar múltiples empresas desde una misma cuenta.

El *panel* de administración organiza las funcionalidades en cuatro grandes secciones, tal como se aprecia en la Figura 5:

- **Compras:** Agrupa las funcionalidades de ingreso de documentos de compra, incluyendo la carga de facturas de compra desde el RCV del SII mediante Rasca, la carga de imágenes de boletas mediante OCR con inteligencia artificial, y la visualización de gastos registrados por imágenes.
- **Ventas:** Concentra las herramientas de registro de ingresos de la empresa, permitiendo la carga de facturas de ventas, boletas electrónicas y ventas procesadas mediante sistemas de punto de venta (POS).
- **Contabilidad:** Integra las funciones del ciclo contable completo, incluyendo el ingreso manual de asientos, la visualización del balance de ocho columnas, el cierre mensual y anual, el Libro de Caja SII y el Libro de Ingresos y Egresos SII.
- **Gestión:** Reúne los módulos de mayor complejidad operacional, entre los que se encuentran la automatización SII para la descarga del RCV mediante Rasca, el módulo de arriendos con extracción de datos por IA, el módulo de Recursos Humanos para la gestión de remuneraciones, el módulo de socios y dividendos, el módulo de patentes municipales, las acciones rápidas ejecutables también desde Renata, y el módulo de solicitudes de clientes para la gestión de *tickets* de soporte.

Esta estructura modular permite al contador acceder de forma inmediata a cualquier funcionalidad sin necesidad de navegar por menús anidados, reduciendo los tiempos de operación y la carga cognitiva asociada al uso del sistema.

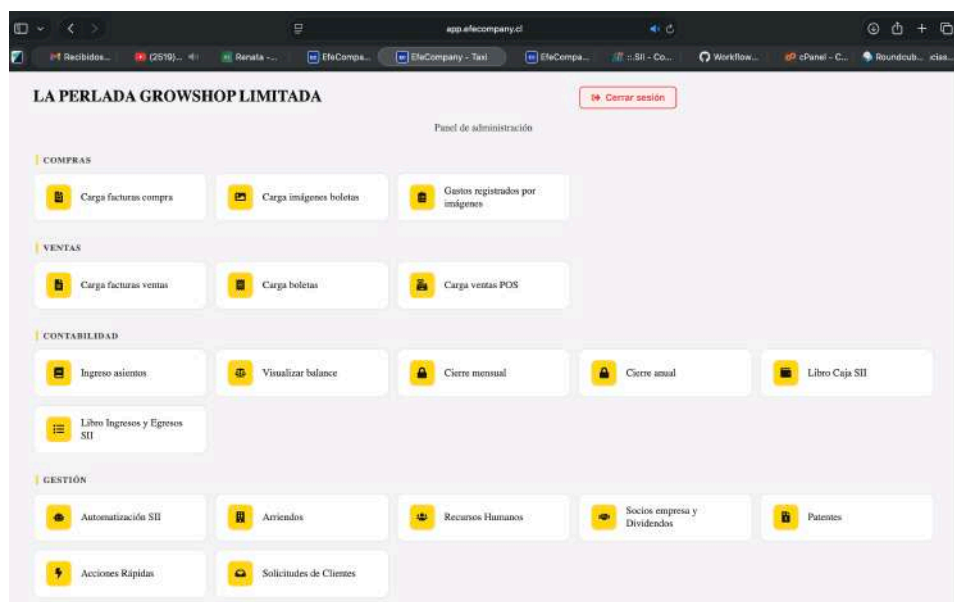


Figura 5 Interfaz de usuario de Taxi

4 Validación de objetivos e hipótesis

4.1 Validación de Objetivos

A continuación se detallan los objetivos iniciales y como estos fueron validados o ejecutados:

- **Implementación de módulos OCR:** Se implementaron extractores de información mediante dos proveedores de inteligencia artificial: el modelo *Claude Sonnet* de Anthropic para documentos en formato PDF y el modelo *GPT-4o-mini* de OpenAI para documentos en formato imagen. Los extractores cubrieron certificados Previred, facturas, boletas, contratos de arriendo, Formulario F-22 y datos de socios. **Resultado:** cumplido.
- **Implementación y evaluación de RAG:** Se implementó un sistema RAG basado en 25 circulares del SII con *embeddings* generados mediante *text-embedding-ada-002* y almacenamiento vectorial en PostgreSQL. La evaluación comparativa frente a un LLM sin RAG evidenció que el modelo RAG obtuvo un F1-score de 0,33 frente a 0,27 del modelo base, mostrando mayor cobertura aunque menor precisión. Sin embargo, las limitaciones ante consultas fuera del dominio indexado y el alto costo de mantención de la base normativa determinaron su descarte del sistema productivo. **Resultado:** implementado, evaluado y descartado por razones técnicas y operacionales documentadas en la sección 3.6.
- **Desarrollo de interfaz de interacción:** Se desarrolló una interfaz web de conversación en tiempo real mediante React y Socket.io, integrada al agente autónomo Renata. Esta interfaz reemplazó el canal WhatsApp previsto inicialmente, aportando además trazabilidad completa de las interacciones para la mejora continua del agente. **Resultado:** cumplido.
- **Evaluación de la reducción de dependencia del conocimiento contable especializado:** Se evaluó en qué medida el sistema permitió a los dueños de empresa ejecutar el proceso de declaración mensual de IVA sin requerir la intervención del contador en la etapa de cálculo. El experimento, descrito en la sección 4.2.3, involucró 25 empresas. En todos los casos, el cálculo del Formulario 29 y la generación del archivo TXT fueron realizados íntegramente por el agente Renata a solicitud del dueño de empresa, sin que el contador participara en ninguna etapa de cálculo ni revisión. La única intervención requerida

fue la carga del archivo TXT al portal del SII y el pago correspondiente, tarea de ejecución mecánica que no exige formación contable y que puede ser realizada por el propio dueño de empresa. **Resultado:** cumplido.

4.2 Validación de Hipótesis

Con el propósito de evaluar la hipótesis planteada desde perspectivas complementarias, se diseñó un esquema de validación cuasi-experimental estructurado en dos experimentos diferenciados según el perfil de usuario y el fenómeno a medir.

- Experimento 1 — Grupo A: Se evaluó el desempeño de la plataforma en el contexto profesional contable, comparando su uso frente a métodos manuales en planillas Excel y softwares comerciales disponibles en el mercado chileno. Para operacionalizar esta evaluación se aplicaron dos instrumentos validados: el cuestionario SUS [11] para medir la usabilidad percibida, y el índice NASA-TLX [12][13][14] para cuantificar la carga cognitiva asociada a cada condición evaluada. El grupo A consiste en 5 Contadores profesionales.
- Experimento 2 — Grupo B: Se evaluó en qué medida el sistema redujo la dependencia del conocimiento contable en el proceso de declaración mensual de IVA, poniendo la plataforma directamente en manos de los dueños de empresa. El grupo B consiste en 25 empresas de distintos rubros y tamaños.

Ambos experimentos se detallan en las secciones siguientes.

4.2.1 Experimento 1 - Grupo A

Objetivo: El objetivo de este experimento consistió en evaluar si la plataforma desarrollada mejoró la productividad, la precisión y la usabilidad en tareas tributarias en comparación con los métodos manuales y los softwares comerciales. Para operacionalizar la variable dependiente general definida en la hipótesis, se establecieron los siguientes indicadores de medición:

- El tiempo de ejecución medido en minutos por cada proceso
- La tasa de error calculada según el porcentaje de discrepancias encontradas
- El nivel de automatización definido por el porcentaje de tareas ejecutadas de forma autónoma por el sistema
- La usabilidad percibida cuantificada mediante la escala SUS [11]
- La carga cognitiva evaluada a través del índice NASA-TLX [12][13][14]

En este experimento, todos los contadores trabajaron con un único *dataset* real correspondiente a una misma empresa asegurando así condiciones homogéneas para la comparación. A partir de este conjunto de datos cada participante ejecutó las mismas actividades estandarizadas (registro de facturas y boletas, generación de balance de ocho columnas y propuesta de declaración de IVA). Con el fin de evaluar diferencias de desempeño según la herramienta utilizada, se solicitó a cada contador repetir el procedimiento en tres escenarios distintos:

- Utilizando exclusivamente planillas de Excel
- Empleando un software contable comercial de su elección
- Utilizando la plataforma desarrollada en el marco de esta tesina

Preparación del entorno de prueba

El grupo de contadores trabajó sobre un mismo conjunto de datos reales, correspondientes a una pequeña empresa chilena del régimen tributario D3. Dicho conjunto incluyó:

- 7 facturas de compras

- 25 facturas de ventas
- 2 boletas de ventas
- 627 registros de ventas con tarjeta
- 10 boletas en formato de imagen
- Información necesaria para generar un balance de ocho columnas

Las pruebas experimentales se llevaron a cabo de manera presencial. El rol del investigador se limitó exclusivamente a entregar instrucciones sobre el uso de los cuestionarios de evaluación SUS, NASA-TLX, y aclarar dudas relacionadas con la forma de responder, sin intervenir en la ejecución de las tareas ni influir en las respuestas. Este enfoque corresponde a la metodología de pruebas acompañadas, ampliamente documentada en la literatura de usabilidad [40]. Dicho acompañamiento permite asegurar la correcta comprensión de los instrumentos y aumentar la fiabilidad de los datos, evitando sesgos derivados de interpretaciones incorrectas.

La elección de NASA-TLX en esta investigación responde a su validez y confiabilidad demostrada en contextos tecnológicos y de interacción humano y computador incluyendo aplicaciones en software interactivo. En el marco de esta tesina su aplicación resultó pertinente para complementar los resultados de usabilidad obtenidos con el SUS, ya que permitió cuantificar el esfuerzo subjetivo requerido por los usuarios al interactuar con el sistema propuesto en comparación con métodos manuales y softwares comerciales. Para una aplicación más relacionada con el contexto chileno se utilizó un documento generado por ISL (Instituto de Seguridad Laboral) del Ministerio del Trabajo y Previsión Social de Chile [12]. A continuación, se revisan los resultados.

Resultados

- a) Tiempo de ejecución: En la Tabla 1 se refleja que los tiempos en el escenario manual requieren una alta demanda temporal. Por su parte el uso de otros ERP redujo la media lo que confirma el impacto positivo de la automatización parcial en la eficiencia de los procesos contables. Finalmente, el uso de la plataforma desarrollada presentó un tiempo inferior, siendo la alternativa más eficiente y consistente lo que se observa en su baja desviación estándar.

Tabla 1. Tiempos totales al realizar todas las tareas y resultados, medidos en minutos

Medida	Manual	Otros ERP	Software Tesina
Media	297,6	69,8	20,8
Desviación Estándar	33,6	12,8	2,8

- b) Tasa de error: Se decidió no considerar esta métrica. Aunque inicialmente los errores parecían medibles de manera uniforme, un análisis cualitativo detallado reveló que su naturaleza difiere sustancialmente según el entorno. En el caso del software desarrollado, los errores se asociaron principalmente a aspectos de interfaz de usuario y experiencia de interacción (confusión por disposición de botones, colores o flujos poco familiares), mientras que en los ERP comerciales los errores estuvieron relacionados con la digitación manual de documentos tributarios, lo cual responde a fuentes cognitivas y procedimentales distintas.
- c) Nivel de automatización: como se visualiza en Tabla 2, los resultados muestran diferencias claras entre las tres condiciones evaluadas. En el método manual no se observa presencia de automatización, lo que obliga a una mayor intervención del usuario en cada tarea. En el caso de los ERP tradicionales se evidencia un nivel bajo de automatización, donde ciertas funciones apoyan al usuario, pero aún persisten procesos dependientes de la digitación y validación manual. En contraste, el software

desarrollado en esta tesina alcanza un nivel alto y consistente de automatización, reduciendo la carga operativa y estandarizando los procedimientos.

Los datos fueron recolectados de los formularios completados por los contadores, en la cual se definieron 5 actividades generales en la cual cada 1 de ellos debían marcar si parte de la actividad requería intervención manual. Respecto a los datos de “Otros ERP” hace referencia a los procesos donde los contadores podían utilizar cualquier otro software disponible que ellos utilicen para realizar las actividades definidas en el formulario, como muestra la Figura 6.

Formulario CuasiExperimental Contadores

Felipe Charme

Fecha 21-08-2025

Minutos totales	porcentaje de errores	Porcentaje de Automatización	ID Participante
303	60%	0%	1

Actividades	Fecha	Hora de Inicio	Hora de fin	Duración en Minutos	hubieron errores	fue necesario hacerlo manual
Ingreso de asientos contable de facturas de compra	21-08-25	13:34	14:30	56	☒	☒
Ingreso de asientos contable de facturas de venta	21-08-25	14:35	17:27	172	☒	☒
Ingreso de asientos contable de boletas de compras	21-08-25	18:00	18:15	15	☒	☒
Generación de balance de 8 columnas	21-08-25	18:20	19:00	40	☐	☒
Generación de Propuesta de declaración de IVA	21-08-25	19:15	19:35	20	☐	☒

Figura 6 Formulario para contadores

Tabla 2. Porcentajes de automatización de experimento 1 – grupo A

Contador	Manual	Otros ERP	Software Tesina
Media	0,0	0,2	0,8
Desviación Estándar	0,00	0,09	0,14

- d) SUS: La Tabla 3 presenta una evaluación y análisis más adecuado de los resultados. Para esto se utilizó un estudio realizado en 2018 por James R. Lewis [41] respecto a la forma de interpretar los puntajes obtenidos. Los resultados obtenidos a partir de la aplicación del SUS muestran diferencias claras en los niveles de usabilidad entre los tres ambientes evaluados. El software desarrollado en la tesina alcanzó un puntaje superior a 90 lo que estaría dentro de los valores más altos posibles en estudios de usabilidad, superando el promedio de referencia de 68 puntos. En contraste, los sistemas ERP comerciales obtuvieron una puntuación intermedia cercana al umbral de aceptabilidad lo que indica una usabilidad promedio. Finalmente, el método manual evidenció un puntaje muy bajo reflejando que este entorno carece de condiciones mínimas para ser considerado un sistema usable en términos de experiencia de usuario.

Tabla 3. Puntajes promedios SUS

Ambiente	Puntuación Promedio	Calificación (Escala Curva)
Software Tesina	92,5	A+
Otros ERP	70	C
Manual	32	F

- e) NASA-TLX: Los resultados obtenidos están representados en la Tabla 4 mostrando diferencias consistentes entre las condiciones evaluadas. En la versión RAW-TLX la modalidad manual presentó la mayor carga de trabajo percibida situándose en un nivel alto, el uso de otros ERP alcanzó un nivel intermedio asociado a una carga baja-media, mientras que el software desarrollado reflejó la menor carga clasificada como baja. Por su parte el análisis del TLX ponderado confirmó esta tendencia lo que respalda la hipótesis de que el software desarrollado logra una reducción significativa en la carga cognitiva y de esfuerzo en comparación con los métodos tradicionales y los sistemas disponibles en el mercado. Un aspecto relevante de los resultados ponderados es que la mayoría de los pesos se concentraron en la dimensión “demanda mental” lo que evidencia que independientemente del ambiente de trabajo, el principal factor en la ejecución de las operaciones contables y tributarias es de carácter cognitivo. En contraste la dimensión “demanda física” fue la menos seleccionada, lo que refuerza que estas tareas se sustentan principalmente en el procesamiento mental y el conocimiento especializado, más que en el esfuerzo físico.

Tabla 4. Puntajes NASA-TLX ponderados

Contador	Manual	Otros ERP	Software Tesina
Media	1343	923	324
Desviación Estándar	170	168	201

4.2.2 Experimento 2 - Grupo B

El segundo experimento se realizó durante la primera semana de marzo de 2026 e involucró a 25 empresas de distintos rubros y regímenes tributarios. El objetivo fue evaluar en qué medida el sistema permitía a los dueños de empresa ejecutar el proceso de declaración mensual de IVA correspondiente al período tributario de febrero de 2026, sin requerir la intermediación técnica de un contador para la etapa de cálculo.

El flujo del experimento fue el siguiente: cada dueño de empresa solicitó a Renata el cálculo de su declaración mensual de IVA. El agente procesó la información tributaria disponible en el sistema, realizó el cálculo del Formulario 29 y generó el archivo TXT correspondiente, entregando al usuario el monto a pagar y el enlace de pago para realizar la transferencia. Posteriormente, el investigador realizó la carga del archivo TXT al portal del SII para cada empresa. Con esta información, los dueños de empresa realizaron el pago directamente al SII, conociendo con precisión el monto y la fecha de vencimiento, sin haber requerido en ningún momento la intervención del contador para calcular, revisar ni explicar los valores.

El 100% de los cálculos fue completado por Renata sin errores, lo que permitió que la totalidad de las 25 empresas participantes completaran su declaración mensual de IVA. Este resultado evidenció que el sistema traslada efectivamente la capacidad de ejecución del proceso tributario hacia el dueño de empresa, reduciendo el rol del contador a una tarea operativa que no exige juicio contable especializado. Los cálculos fueron comparados con los cálculos manuales, con el fin de validar los cálculos adecuados.

Es importante indicar que se identificaron las siguientes características que compartían las empresas, estas fueron de regímenes 14D3, 14D8, 14A, sin empleados y sin compras internacionales. Ya que este tipo de empresas no cuentan con toda la información necesaria y adecuada dentro del portal del SII.

4.2.3 Experimento 3 - modelo RAG

Este enfoque entrega lineamientos claros respecto a la necesidad de establecer métricas de exactitud y procedimientos de validación controlados, con el fin de reducir sesgos y asegurar la replicabilidad de los resultados. Para esto se utilizó la metodología de S. Simon et al [15]. No obstante, la metodología no se aplicó de manera literal, ya que el objeto de estudio de esta investigación corresponde a un único modelo RAG especializado en normativa tributaria chilena y documentos empresariales. En consecuencia, se adaptó los principios generales del marco tales como la definición de un banco de preguntas de referencia, y la comparación contra un *gold standard* experto y un conjunto de datos propios específicos de la realidad tributaria local. Esta adaptación permitió mantener la rigurosidad metodológica propuesta por la literatura, pero aseguró al mismo tiempo la pertinencia del experimento en relación con los objetivos de la tesina. Para la evaluación se aplicaron las métricas definidas en la sección 2.3 (Fórmulas 1, 2 y 3)

Objetivo: Cuantificar si el *pipeline RAG* (*embeddings* + recuperación + generación) alcanza y supera un umbral de exactitud aceptable en consultas tributarias chilenas, y si mejora significativamente respecto de un LLM vanilla sin recuperación externa.

Se utilizó un banco de preguntas con respuestas estándar derivadas de documentos tributarios oficiales, y se compararon dos condiciones, un modelo LLM de última generación (ChatGPT-5) y un modelo RAG construido con *embeddings* y recuperación en base a las circulares del SII [39]. Este enfoque mantiene la lógica metodológica del *blueprint* propuesto que es uso de *baseline*, métricas de precisión, *recall* y F1-score.

Variables dependientes: Precision, *Recall*, F-1

Variables Independientes: Arquitectura RAG del trabajo, *ChatGPT* sin *RAG*

Los resultados obtenidos se representan en la Tabla 5, que reflejan diferencias relevantes entre el desempeño de GPT-5 y el modelo RAG del trabajo en la tarea de responder preguntas normativas tributarias. GPT-5 alcanzó una mayor precisión lo que indica que cuando proporciona una respuesta esta es más frecuentemente correcta respecto al estándar de referencia. Sin embargo, este modelo presentó un *recall* bajo, lo que significa que dejó sin responder una proporción importante de preguntas cuya respuesta sí estaba contenida en la base normativa.

Tabla 5. Resultados evaluación RAG vs GPT 5

modelo	TP	FP	FN	Precisión	Recall	F1
GPT-5	8	4	38	0,6667	0,1739	0,2759
RAG	10	13	27	0,4348	0,2703	0,3333

Por el contrario, el modelo RAG mostró un comportamiento distinto, aunque su precisión fue menor es decir incurrió en más respuestas incorrectas, logró un *recall* superior, lo que sugiere que recuperó una mayor cantidad de respuestas existentes en el corpus. Como consecuencia de este equilibrio el RAG alcanzó un F1-score mayor que GPT-5 lo que evidencia que pese a ser menos preciso su capacidad de cubrir preguntas del dominio normativo fue más robusta. En síntesis, GPT-5 se caracteriza por priorizar la exactitud en las respuestas, mientras que el modelo RAG privilegia la cobertura, aunque a costa de introducir errores adicionales.



5 Conclusiones

El presente trabajo abordó el diseño e implementación de un sistema integrado compuesto por tres componentes principales: el ERP Taxi, el agente de inteligencia artificial Renata y el *scraper* tributario Rasca. El objetivo central fue evaluar en qué medida la automatización del cálculo tributario reduce la dependencia de los dueños de pequeñas empresas chilenas respecto del conocimiento contable especializado para la ejecución de procesos tributarios periódicos. Este resultado constituyó evidencia concreta de que el sistema reduce la participación del conocimiento contable especializado en el proceso de declaración mensual de IVA, trasladando la capacidad de ejecución hacia el dueño de empresa. No obstante, es importante precisar que el alcance de esta validación se circunscribe a la declaración de IVA como proceso tributario específico, y que el proceso no está completamente automatizado, dado que la carga del archivo y el pago al SII siguen requiriendo intervención humana.

Se concluye que la integración de un ERP con un agente de inteligencia artificial y automatización por OCR constituye una solución técnicamente viable para reducir la dependencia de los dueños de pequeñas empresas chilenas respecto del conocimiento contable especializado en la ejecución de procesos tributarios periódicos, aportando mayor autonomía operacional y reduciendo el rol del contador hacia tareas de ejecución que no requieran juicio técnico especializado.

En relación con la seguridad y la gestión de la información, se estima que los trabajos futuros deben priorizar el fortalecimiento de los protocolos de ciberseguridad para asegurar la integridad y la disponibilidad de los datos contables y financieros procesados por la plataforma. Se planteó la necesidad de implementar mecanismos avanzados de gobernanza de datos y políticas de privacidad que garantizaran la protección de la información sensible frente a posibles amenazas externas.



Referencias

- [1] Servicio de Impuestos Internos. (2024). Estadísticas de empresas por actividad económica [Online]. Available: https://www.sii.cl/sobre_el_sii/estadisticas_de_empresas.html
- [2] A. M. Vargas Cárdenas y J. Trujillo Puentes. (2021). La tributación de las PYMES y su cumplimiento tributario [Online]. Available: <https://cet.uchile.cl>
- [3] Nubox. (2024). Automatización contable para PYMES [Online]. Available: <https://www.nubox.com>
- [4] Rindegastos. (2024). Plataforma de rendición de gastos para empresas [Online]. Available: <https://www.rindegastos.com>
- [5] Servicio de Impuestos Internos. (2024). Tipos de Regímenes Tributarios [Online]. Available: https://www.sii.cl/destacados/modernizacion/tipos_regimenes_mt.html
- [6] Servicio de Impuestos Internos. (2024). Preguntas Frecuentes - Declaración mensual de IVA [Online]. Available: https://www.sii.cl/preguntas_frecuentes/impuestos_mensuales/001_130_1726.htm
- [7] Defontana. (2024). ERP Pymes Génesis [Online]. Available: <https://www.defontana.com/cl/erp-genesis>
- [8] Databot. (2024). La nueva era de la IA en servicio al cliente [Online]. Available: <https://www.databot.cl>
- [9] Veryfi. (2024). WhatsApp Document Extraction and OCR [Online]. Available: <https://www.veryfi.com/products/whatsapp-document-extraction-ocr/>
- [10] QuickBooks. (2024). Software de Contabilidad Online [Online]. Available: <https://quickbooks.intuit.com>
- [11] J. Brooke. (1996). SUS: A quick and dirty usability scale [Online]. Available: <https://hell.meiert.org/core/pdf/sus.pdf>
- [12] Instituto de Seguridad Laboral. (2024). Manual de aplicación NASA-TLX [Online]. Available: https://ergomedia.isl.gob.cl/app_ergo/nasatlx/nasa-tlx.pdf
- [13] NASA Ames Research Center. (2024). Task Load Index (NASA-TLX) [Online]. Available: <https://humansystems.arc.nasa.gov/groups/TLX/tlxpaperpencil.php>
- [14] S. G. Hart y L. E. Staveland. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research [Online]. Available: [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- [15] S. Simon, A. Mailach, J. Dorn y N. Siegmund. (2024, Octubre). A methodology for evaluating RAG systems: A case study on configuration dependency validation [Online]. Available: <https://arxiv.org/abs/2410.08801>
- [16] J. A. Romero, A. la Paz y M. J. García. (2024, Abril). Impacto de las herramientas tecnológicas en las ventas de empresas chilenas [Online]. Available: https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-07642024000200011
- [17] W. S. McCulloch y W. A. Pitts. (1943). A logical calculus of the ideas immanent in nervous activity [Online]. Available: <https://doi.org/10.1007/BF02478259>
- [18] McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955). A proposal for the Dartmouth summer research project on artificial intelligence.
- [19] Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6), 386.
- [20] Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433-460.
- [21] J. McCarthy, "What is Artificial Intelligence?," Computer Science Department, Stanford University, Stanford, CA, 2007. [Online]. Available: <http://jmc.stanford.edu/artificial-intelligence/what-is-ai.html>
- [22] J. D. Kelleher, Deep Learning. Cambridge, MA, USA: MIT Press, 2019.
- [23] T. M. Mitchell, Machine Learning. New York, NY, USA: McGraw-Hill, 1997.
- [24] Y. LeCun, Y. Bengio y G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-444, 2015
- [25] OpenAI. (2024). ChatGPT API Documentation [Online]. Available: <https://openai.com/es-419/index/chatgpt/>
- [26] J. R. Casar Corredera, "IA generativa," *Anales de la Real Academia de Doctores de España*, vol. 8, no. 3, pp. 475-489, 2023.
- [27] K. T. Chitty-Venkata, S. Mittal, M. Emani, V. Vishwanath, y A. K. Somani, "A survey of techniques for optimizing transformer inference," *Journal of Systems Architecture*, vol. 144, p. 102990, 2023
- [28] W. Li, W. Liu, M. Deng, X. Liu, and L. Feng, "The impact of large language models on accounting and future application scenarios," *Journal of Accounting Literature*, 2025. doi: 10.1108/JAL-12-2024-0357.
- [29] K. W. Church et al., "Emerging trends: a gentle introduction to RAG," *Natural Language Engineering*, vol. 30, no. 4, pp. 870-881, 2024. doi:10.1017/S1351324924000044
- [30] S. Wu et al., "Retrieval-Augmented Generation for Natural Language Processing: A Survey," *arXiv*, 2025. [Online]. Disponible en: <https://arxiv.org/abs/2407.13193>.
- [31] L. Wang et al., "A Survey on Large Language Model based Autonomous Agents," *Frontiers of Computer Science*, pp. 1-42, 2025, doi: 10.1007/s11704-024-40231-1.



- [32] J. Wang, J. Tang, M. Yang, X. Bai, y J. Luo, "Improving OCR-Based Image Captioning by Incorporating Geometrical Relationship", en Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1306-1315, 2021
- [33] G. Me, C. Calero and P. Lago, "Architectural Patterns and Quality Attributes Interaction," 2016 Qualitative Reasoning about Software Architectures (QRASA), Venice, Italy, 2016, pp. 27-36, doi: 10.1109/QRASA.2016.10.
- [34] F. J. Llamas Fernández y J. C. Fernández Rodríguez, "La metodología Lean Startup: desarrollo y aplicación para el emprendimiento," Revista EAN, vol. 84, pp. 79–95, 2018, doi: 10.21158/01208160.n84.2018.1918
- [35] D. Cantone, Implementación y Debugging. Buenos Aires, Argentina: Usershop, 2006.
- [36] N. A. Azman, A. Mohamed, and A. M. Jamil, "Artificial Intelligence in Automated Bookkeeping: A Value-added Function for Small and Medium Enterprises," Journal of Information and Visualization, vol. 5, no. 3, 2021. doi: 10.30630/joiv.5.3.669.
- [37] J. Sepúlveda Rivera, «IA. Desafíos para las mujeres contables», Summa. arith., vol. 1, n.º 10, pp. 59–64, ene. 2025.
- [38] M. Cartes López, "Propuesta metodológica para diseñar un modelo de Auditoría Predictiva," Universidad de Chile, 2024. [En línea]. Disponible en: <https://repositorio.uchile.cl/handle/2250/206356>
- [39] Servicio de Impuestos Internos. (2025). Circulares SII año 2025 [Online]. Available: https://www.sii.cl/normativa_legislacion/circulares/2025/indcir2025.html
- [40] J. Nielsen, Usability Engineering. San Diego, CA, USA: Academic Press, 1994.
- [41] J. R. Lewis and J. Sauro, "Item benchmarks for the System Usability Scale," Journal of Usability Studies, vol. 13, no. 3, pp. 158–167, May 2018.
- [42] Anthropic. (2025). Model Context Protocol Architecture [Online]. Available: <https://modelcontextprotocol.io/docs/learn/architecture>



Anexos

I. Documento de consentimiento para contadores

Consentimiento Informado para Participación en Tesina

Título de la Tesina: Diseño e Implementación de un ERP con Automatización por Inteligencia Artificial y OCR para Aumentar la Independencia Tributaria de Pequeñas Empresas Chilenas.

Investigador Responsable:
Fernolds Alvarado
Magíster en Tecnologías de la Información
Universidad Técnica Federico Santa María

Compromiso de Confidencialidad

Los datos utilizados durante la investigación serán confidenciales. En la tesina pública no se incluirán nombres, RUT ni información identificable de los contadores o de las empresas con las que trabajan. Los resultados se presentarán de forma agregada y general. Una bitácora privada con información detallada será mantenida solo para revisión académica en caso de ser requerida.

Consentimiento

Declaro que he sido informado/a del propósito de este estudio, de los datos que serán recolectados y de su uso académico. Autorizo mi participación en este estudio, entendiendo que los resultados de la tesina serán de carácter público, manteniendo siempre la confidencialidad de mis datos personales y profesionales.

Nombre del Contador/a: _____

RUT: _____

Firma: _____

Fecha: ____/____/____

II. Documento de consentimiento para dueños de negocios

Consentimiento Informado para Participación en Tesina

Título de la Tesina: Diseño e Implementación de un ERP con Automatización por Inteligencia Artificial y OCR para Aumentar la Independencia Tributaria de Pequeñas Empresas Chilenas.



Investigador Responsable:
Fernolds Alvarado
Magíster en Tecnologías de la Información
Universidad Técnica Federico Santa María

Compromiso de Confidencialidad

Los datos contables y administrativos que se utilicen serán para temas académicos. Una bitácora privada con la información será resguardada de manera confidencial únicamente para fines de validación por parte de los evaluadores académicos.

Consentimiento

Declaro que he sido informado/a del propósito de esta investigación, de los datos que serán recolectados y de su uso académico. Autorizo la participación de mi empresa en este estudio, entendiendo que los resultados de la tesina serán de carácter público, manteniendo siempre la confidencialidad de los datos sensibles.

Nombre de la Empresa: _____

RUT de la Empresa: _____

Nombre Representante Legal _____

RUT Representante Legal _____

Firma: _____

Fecha: ____/____/____

III. Encuesta SUS

Preguntas	Totalmente en desacuerdo	En desacuerdo	Neutro	De acuerdo	Totalmente de acuerdo
Creo que me gustaría utilizar este sistema con frecuencia	FALSO	FALSO	FALSO	FALSO	FALSO
Encontré el sistema innecesariamente complejo	FALSO	FALSO	FALSO	FALSO	FALSO
Pensé que el sistema era fácil de usar	FALSO	FALSO	FALSO	FALSO	FALSO
Creo que necesitaría el apoyo de un técnico para poder utilizar este sistema	FALSO	FALSO	FALSO	FALSO	FALSO



Encontré que las diversas funciones de este sistema estaban bien integradas	FALSO	FALSO	FALSO	FALSO	FALSO
Pensé que había demasiada inconsistencia en este sistema	FALSO	FALSO	FALSO	FALSO	FALSO
Me imagino que la mayoría de la gente aprendería a utilizar este sistema muy rápidamente	FALSO	FALSO	FALSO	FALSO	FALSO
Encontré el sistema muy complicado de usar	FALSO	FALSO	FALSO	FALSO	FALSO
Me sentí muy seguro usando el sistema	FALSO	FALSO	FALSO	FALSO	FALSO
Necesitaba aprender muchas cosas antes de empezar con este sistema	FALSO	FALSO	FALSO	FALSO	FALSO

IV. Curved Grading Scale for the SUS

Grade	SUS	Percentile range
A +	84.1 - 100	96 - 100
A	80.8 - 84.0	90 - 95
A-	78.9 - 80.7	85 - 89
B+	77.2 - 78.8	80 - 84
B	74.1 - 77.1	70-79
B-	72.6 -74.0	65-69
C +	71.1 - 72,5	60 - 64
C	65.0 - 71.0	41 - 59
C-	62.7 - 64.9	35 - 40
D	51.7 - 62.6	15 - 34
F	0- 51.6	0 -14

V. Encuesta NASA -TLX con tareas evaluadas y las 6 dimensiones evaluadas

M F T R E Fr

	¿Cuánta actividad mental y perceptiva fue necesaria? (pensar, decidir, calcular, etc.). ¿Es una tarea difícil o fácil, simple o compleja, pesada o ligera?	¿Cuánta actividad física fue necesaria? ¿Se trata de una tarea difícil o fácil, lenta o rápida, relajada o cansada?	¿Cuánta presión de tiempo sintió debido al ritmo al cual se sucedían las tareas o elementos de las tareas? ¿Era el ritmo lento y pausado, o rápido y frenético?	¿Hasta qué punto cree que ha tenido éxito en los objetivos establecidos por el investigador (o por Ud. mismo)? ¿Cuál es su grado de satisfacción con el nivel de ejecución?	¿En qué medida ha tenido que trabajar (física o mentalmente) para alcanzar su nivel de resultados?	Durante la tarea, ¿en qué medida se ha sentido inseguro, desalentado, irritado, tenso o preocupado o por el contrario, se ha sentido seguro, contento, relajado y satisfecho?	
Condición	Demanda mental (0–100)	Demanda física (0–100)	Demanda temporal (0–100)	Rendimiento (0=bueno, 100=pobre)	Esfuerzo (0–100)	Frustración (0–100)	Raw TLX (promedio simple)
Ingreso de asientos contable de facturas de compra							
Ingreso de asientos contable de facturas de venta							
Ingreso de asientos contable de boletas de compras							
Ingreso de asientos contable de boletas de ventas							
Generacion de balance de 8 columnas							
Generacion de Propuesta de declaracion de IVA							
Promedio							

VI. Tabla de elección para cálculo de peso de priorización de dominios



M	F	F	T	T	E
FALSO	FALSO	FALSO	FALSO	FALSO	FALSO
M	T	F	R	T	Fr
FALSO	FALSO	FALSO	FALSO	FALSO	FALSO
M	R	F	E	R	E
FALSO	FALSO	FALSO	FALSO	FALSO	FALSO
M	E	F	Fr	R	Fr
FALSO	FALSO	FALSO	FALSO	FALSO	FALSO
M	Fr	T	R	E	Fr
FALSO	FALSO	FALSO	FALSO	FALSO	FALSO

VII. Cálculo de puntuación convertida

Variable	(a) Peso	(b) Puntuación	(c) Puntuación convertida (b x a)
Exigencias Mentales			
Exigencias Físicas			
Exigencias Temporales			
Rendimiento			
Esfuerzo			
Frustración			
TOTAL			

VIII. criterio de interpretación de puntuaciones

NASA TLX	Nivel de Carga Mental
500 puntos o menos	Bajo
Sobre los 500 puntos y por debajo de los 1000 puntos	Medio



Evaluación global sobre 1000 puntos y cuyos factores predominantes posean un carácter intrínseco (independientemente que la suma de factores intrínsecos no alcance los 100 puntos)	Alto
---	------

IX. Consulta para buscar vector más cercano

```
const sql = `
SELECT contenido
FROM public.documentos_vectores
ORDER BY embedding <=> $1::vector
LIMIT 1;
`;
```

X. Formato Encuesta

Formulario CuasiExperimental Contadores

Nombre Contador

Fecha

Minutos totales	porcentaje de errores	Porcentaje de Automatización
00:00:00	0%	100%

ID Participante

1

Actividades	Fecha	Hora de Inicio	Hora de fin	Duración en Minutos	hubieros errores	fue necesario hacerlo manual
Ingreso de asientos contable de facturas de compra				00:00:00	☐	☐
Ingreso de asientos contable de facturas de venta				00:00:00	☐	☐
Ingreso de asientos contable de boletas de compras				00:00:00	☐	☐
Generacion de balance de 8 columnas				00:00:00	☐	☐
Generacion de Propuesta de declaracion de IVA				00:00:00	☐	☐

Manual - ExcelSUS- ManualNASA-TLX-ManualOtros SistemasSUS- otrosNASA-TLX-otrosTaxiSUS-TAXINASA-TLX-TAXI+