



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Facultad de Ingeniería de Sistemas e Informática

Escuela Profesional de Ingeniería de Sistemas

**Reingeniería del sistema web SIE para mejorar la
gestión de expedientes judiciales en la Procuraduría
Pública del Sector Interior – MININTER**

TESIS

Para optar el Título Profesional de Ingeniero de Sistemas

AUTOR

Kevin Miguel ORTIZ ABANTO

Nabia Jasmin PACHAS LOPEZ

ASESOR

Dr. Frank Edmundo ESCOBEDO BAILON

Lima, Perú

2026



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Ortiz, K. & Pachas, N. (2026). *Reingeniería del sistema web SIE para mejorar la gestión de expedientes judiciales en la Procuraduría Pública del Sector Interior – MININTER*. [Tesis de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ingeniería de Sistemas e Informática, Escuela Profesional de Ingeniería de Sistemas]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios autor/ asesor

Datos de autor	
Nombres y apellidos	Nabia Jasmin Pachas Lopez
Tipo de documento de identidad	DNI
Número de documento de identidad	72606359
URL de ORCID	https://orcid.org/0009-0005-8619-097X
Datos de autor	
Nombres y apellidos	Kevin Miguel Ortiz Abanto
Tipo de documento de identidad	DNI
Número de documento de identidad	72141765
URL de ORCID	https://orcid.org/0009-0006-0202-3400
Datos de asesor	
Nombres y apellidos	Frank Edmundo Escobedo Bailon
Tipo de documento de identidad	DNI
Número de documento de identidad	41671087
URL de ORCID	https://orcid.org/0000-0002-2058-0976
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Cesar Alcántara Loayza
Tipo de documento	DNI
Número de documento de identidad	09132297
Miembro del jurado 1	
Nombres y apellidos	Ana Maria Huayna Dueñas
Tipo de documento	DNI
Número de documento de identidad	06017183
Datos de investigación	

Línea de investigación	Ingeniería de software
Grupo de investigación	Ingeniería web
Agencia de financiamiento	Ninguna (Fondos propios)
Ubicación geográfica de la investigación	Lima
Año o rango de años en que se realizó la investigación	2024 - 2025
URL de disciplinas OCDE	https://purl.org/pe-repo/ocde/ford#2.02.04



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
FACULTAD DE INGENIERIA DE SISTEMAS E INFORMATICA
Escuela Profesional de Ingeniería de Sistemas

Acta Virtual de Sustentación de Tesis

Siendo las 15:00 horas del día 11 de mayo del año 2026, se reunieron virtualmente los docentes designados como miembros de Jurado de Tesis, presidido por el Mg. Cesar A. Alcántara Loayza, Mg. Ana Maria Huayna Dueñas (Miembro) y el Dr. Frank E. Escobedo Bailón (Miembro Asesor), usando la plataforma Meet (<https://meet.google.com/qnb-butc-tcz>) para la sustentación Virtual de la tesis Intitulada: **“REINGENIERÍA DEL SISTEMA WEB SIE PARA MEJORAR LA GESTIÓN DE EXPEDIENTES JUDICIALES EN LA PROCURADURÍA PÚBLICA DEL SECTOR INTERIOR – MININTER.”**, de la bachiller : **PACHAS LOPEZ, Nabia Jasmin**; para obtener el Título Profesional de Ingeniero de Sistemas.

Acto seguido de la exposición de la Tesis, el Presidente invitó a la Bachiller a responder las preguntas formuladas por los Miembros del Jurado.

La Bachiller, en el curso de sus intervenciones demostró pleno dominio del tema, al responder con acierto y fluidez las preguntas formuladas por los señores miembros del Jurado.

Finalmente habiéndose efectuado la calificación correspondiente por los miembros del Jurado, la bachiller obtuvo la nota de 17 (**Diesisiete**)

A continuación, el Presidente del Jurado Mg. Cesar A. Alcántara Loayza, declara a la bachiller **Ingeniero de Sistemas**.

Siendo 16:12: horas, se levantó la sesión.

Mg. Cesar A. Alcántara Loayza
Presidente

Mg. Ana Maria Huayna Dueñas
Miembro

Dr. Frank E. Escobedo Bailón
Miembro Asesor



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS
FACULTAD DE INGENIERIA DE SISTEMAS E INFORMATICA
Escuela Profesional de Ingeniería de Sistemas

Acta Virtual de Sustentación de Tesis

Siendo las 15:00 horas del día 11 de mayo del año 2026, se reunieron virtualmente los docentes designados como miembros de Jurado de Tesis, presidido por el Mg. Cesar A. Alcántara Loayza, Mg. Ana Maria Huayna Dueñas (Miembro) y el Dr. Frank E. Escobedo Bailón (Miembro Asesor), usando la plataforma Meet (<https://meet.google.com/qnb-butc-tcz>) para la sustentación Virtual de la tesis Intitulada: **“REINGENIERÍA DEL SISTEMA WEB SIE PARA MEJORAR LA GESTIÓN DE EXPEDIENTES JUDICIALES EN LA PROCURADURÍA PÚBLICA DEL SECTOR INTERIOR – MININTER.”**, del bachiller : **ORTIZ ABANTO, Kevin Miguel**; para obtener el Título Profesional de Ingeniero de Sistemas.

Acto seguido de la exposición de la Tesis, el Presidente invitó del Bachiller a responder las preguntas formuladas por los Miembros del Jurado.

El Bachiller, en el curso de sus intervenciones demostró pleno dominio del tema, al responder con acierto y fluidez las preguntas formuladas por los señores miembros del Jurado.

Finalmente habiéndose efectuado la calificación correspondiente por los miembros del Jurado, el bachiller obtuvo la nota de 17 **(Diecisiete)**

A continuación, el Presidente del Jurado Mg. Cesar A. Alcántara Loayza, declara a la bachiller **Ingeniero de Sistemas**.

Siendo 16:12: horas, se levantó la sesión.

Mg. Cesar A. Alcántara Loayza
Presidente

Mg. Ana María Huayna Dueñas
Miembro

Dr. Frank E. Escobedo Bailón
Miembro Asesor

Anexo 1
CERTIFICADO DE SIMILITUD

Yo, **FRANK EDMUNDO ESCOBEDO BAILON**, en mi condición de asesor acreditado con Resolución Directoral de la EP N°000012-2025-EPISI-FISI/UNMSM de la **tesis** cuyo título es **REINGENIERÍA DEL SISTEMA WEB SIE PARA MEJORAR LA GESTIÓN DE EXPEDIENTES JUDICIALES EN LA PROCURADURÍA PÚBLICA DEL SECTOR INTERIOR – MININTER.**, presentado por los **bachilleres ORTIZ ABANTO, Kevin Miguel - PACHAS LOPEZ, Nabia Jasmin** para optar el **TITULO PROFESIONAL DE INGENIERO DE SISTEMAS**, CERTIFICO que se ha cumplido con lo establecido en la Directiva de Originalidad y de Similitud de Trabajos Académicos, de Investigación y Producción Intelectual. Según la revisión, análisis y evaluación mediante la plataforma PIDS, el documento evaluado cuenta con el **7%** de similitud, nivel PERMITIDO para continuar con los trámites correspondientes y para su **publicación en el repositorio institucional.**

Se emite el presente certificado en cumplimiento de lo establecido en las normas vigentes, como uno de los requisitos para la obtención del **TITULO PROFESIONAL DE INGENIERO DE SISTEMAS.**

Lima, 9 de junio del 2026

Firma del Asesor:



Nombres y Apellidos del asesor: Frank Edmundo Escobedo Bailon

DNI: 41671087

* Según corresponda: bachiller, magíster, doctor; título profesional o título de segunda especialidad

La vigencia del presente certificado corresponde a un periodo de noventa (90) días desde la fecha de su emisión; siempre que no haya observaciones en el texto durante el proceso de revisión por el jurado informante o surja alguna contingencia.

Agradecimiento

Primero que nada agradecemos a Dios por darnos la oportunidad de vivir esta experiencia única que llamamos vida. A nuestros familiares, quienes nos ayudaron a ser constantes y con los que compartimos momentos que nunca se olvidan. A nuestros docentes, especialmente al Dr. Frank Escobedo Bailón, quien nos guió en este camino con paciencia y más allá de ser nuestro docente fue también un mentor, un amigo, enseñándonos no solo sus conocimientos sino también valores que tendremos presentes para toda la vida. Por último, agradecemos a esas personas que empezaron siendo compañeros de estudio y terminaron por ser más que compañeros, verdaderos amigos. Sin importar la distancia, a ustedes, Rubén, Enrique, Mauro y Lucía, gracias por hacer que cada día fuera especial, que los momentos difíciles y de estrés se conviertan en diversión y risas, y por recordarnos siempre que, con esfuerzo y actitud positiva, todo es posible.

Dedicatoria

Dedicamos esta tesis a nuestros padres por su amor infinito y su apoyo constante, porque sin ellos este logro no hubiera sido remotamente posible. A nuestros profesores y mentores, por su orientación y conocimiento, que nos inspiraron a seguir adelante en nuestro camino universitario. A nuestros amigos Ruben, Mauro y Enrique que nos enseñaron que lo más importante es disfrutar el proceso y que gracias a sus risas, ocurrencias y palabras de aliento hicieron este viaje mucho más llevadero y memorable.

Resumen

La presente investigación tuvo como objetivo hacer una reingeniería el sistema web SIE, con la finalidad de mejorar la gestión de expedientes judiciales en la Procuraduría Pública del Sector Interior del Ministerio del Interior (MININTER). La solución propuesta se organizó en cuatro módulos mejorados: (i) recepción y consulta de expedientes judiciales, mediante el uso de tecnologías de reconocimiento óptico de caracteres (OCR) y modelos de lenguaje de gran escala (LLM), (ii) planificación y gestión de audiencias, (iii) elaboración y envío de recursos legales y (iv) seguimiento de expedientes. El estudio corresponde a una investigación aplicada, con enfoque cuantitativo y diseño cuasiexperimental, considerando dos poblaciones: abogados y expedientes. En el caso de los abogados, se empleó un diseño pretest y posttest, aplicados antes y después de la implementación del sistema mejorado. Para los expedientes, se recopilaron indicadores temporales mediante un grupo de control y un grupo experimental, evaluados antes y después de la implementación del sistema. Los instrumentos de recolección de datos fueron una encuesta de 16 preguntas aplicada a 38 abogados y una ficha de registro utilizada en 496 expedientes. La confiabilidad de la encuesta se evaluó mediante el coeficiente alfa de Cronbach. Asimismo, se aplicaron las pruebas de normalidad de Shapiro-Wilk y Kolmogorov-Smirnov para las poblaciones 1 y 2, respectivamente. Para el análisis de resultados, se empleó la prueba t de muestras relacionadas y la prueba de Wilcoxon en la población de abogados y la prueba U de Mann-Whitney en la población de expedientes. Los resultados evidenciaron mejoras significativas tras la modernización del sistema SIE, con un tamaño de efecto grande (d de Cohen = 2.06), lo que confirma el impacto positivo de la mejora propuesta. Asimismo, el tiempo del proceso de recepción de expedientes se redujo de 14,0 a 7,3 minutos, mientras que el proceso de elaboración de recursos legales disminuyó de 9,5 a 3,5 minutos. En conclusión, la modernización del sistema web SIE permitió mejorar significativamente la gestión de expedientes judiciales en la Procuraduría Pública del Sector Interior, contribuyendo a una mayor eficiencia operativa y optimización de los procesos institucionales.

Palabras Claves: gestión de expedientes - reingeniería de software –LLM – OCR – PNL

Abstract

This research aimed to modernize the SIE web system to improve the management of court files at the Public Prosecutor's Office of the Interior Sector of the Ministry of the Interior (MININTER). The proposed solution was organized into four improved modules: (i) reception and consultation of court files using optical character recognition (OCR) and large-scale language models (LLM), (ii) hearing planning and management, (iii) preparation and submission of legal appeals, and (iv) case tracking. The study is an applied research project with a quantitative approach and a quasi-experimental design, considering two populations: lawyers and case files. For the lawyers, a pretest-posttest design was used, administered before and after the implementation of the improved system. For the case files, time-series indicators were collected using a control group and an experimental group, evaluated before and after the system implementation. The data collection instruments were a 16-question survey administered to 38 lawyers and a registration form used for 496 case files. The reliability of the survey was assessed using Cronbach's alpha coefficient. The Shapiro-Wilk and Kolmogorov-Smirnov normality tests were also applied to populations 1 and 2, respectively. For the analysis of results, the paired-samples t-test and the Wilcoxon signed-rank test were used for the lawyer population, and the Mann-Whitney U test was used for the case file population. The results showed significant improvements after the modernization of the SIE system, with a large effect size (Cohen's $d = 2.06$), confirming the positive impact of the proposed a. Furthermore, the case file reception process time was reduced from 14.0 to 7.3 minutes, while the legal resource preparation process decreased from 9.5 to 3.5 minutes. In conclusion, the modernization of the SIE web system significantly improved the management of judicial files in the Public Prosecutor's Office of the Interior Sector, contributing to greater operational efficiency and optimization of institutional processes.

Keywords: case management - software reengineering - LLM - OCR - NLP

Índice

Planteamiento del Problema	13
1.1. Descripción del Problema	13
1.2. Formulación del Problema	15
1.2.1. Problema General	15
1.2.2. Problemas Específicos	15
1.3. Objetivos de la Investigación	15
1.3.1. Objetivo General	15
1.3.2. Objetivos Específicos	16
1.4. Delimitación de la Investigación	16
1.4.1. Delimitación Conceptual	16
1.4.2. Delimitación Temporal	18
1.4.3. Delimitación Espacial	18
1.5. Justificación de la Investigación	18
1.5.1. Justificación teórica	18
1.5.2. Justificación práctica	19
1.5.3. Justificación metodológica	20
1.6. Hipótesis	20
1.6.1. Hipótesis General	20
1.6.2. Hipótesis Específicas	20
Marco Teórico	21
2.1. Bases Teóricas	21
2.1.1. Reingeniería y modernización de software	21
2.1.1.1. Modelo General de la Reingeniería de Software	21
2.1.2. Herramientas para la modernización de software	22
2.1.2.1. Inteligencia Artificial	22
2.1.2.2. Reconocimiento Óptico de Caracteres	23
2.1.2.3. Procesamiento de Lenguaje Natural	26
2.1.2.4. LLMs en el procesamiento de lenguaje natural	28
2.1.3. Metodología para implementación de LLMs	31
2.1.3.1. Metodología GenAIOps	31
2.1.3.2. Etapas de la metodología GenAIOps	32
2.1.4. Gestión de expedientes judiciales en la procuraduría pública del Sector Interior - MININTER	34
2.1.4.1. Procuradurías públicas del Ministerio del Interior	34
2.1.4.2. Procuraduría Pública del Sector Interior y sus funciones	35
2.1.4.3. Funcionamiento interno de la Procuraduría del Sector Interior.	36
2.1.4.4. Rol de la procuraduría en los procesos penales	37
2.1.4.5. Etapas principales del proceso penal	38
2.1.4.6. Sistemas informáticos utilizados por la Procuraduría Pública del Sector Interior	39
2.1.4.7. Gestión de Expedientes Judiciales en el MININTER.	41

2.1.4.8. Informática jurídica	45
2.1.4.9. E-Justicia	47
2.1.4.10. Aplicación de la IA en el ámbito judicial	48
2.2. Antecedentes	49
2.2.1. Antecedentes Internacionales	49
2.2.2. Antecedentes Nacionales	53
2.3. Marco Conceptual	54
Estado del Arte	58
3.1. Sistemas o Aplicaciones Comerciales	58
3.1.1. Sistema de gestión judicial con inteligencia artificial Onesait Government Justice en la Corte de Cajamarca	58
3.1.2. Sistema de seguimiento y centralización de expedientes judiciales CaseTracking	59
3.1.3. Sistema de gestión judicial integral Atlanta Web para la digitalización del proceso jurisdiccional	59
3.1.4. Sistema de gestión de casos judiciales para la administración integral del ciclo de vida procesal	60
3.1.5. Sistemas de Gestión Procesal con capacidades de automatización avanzada en la Administración de Justicia	60
3.1.6. Sistema de gestión procesal Avantius para la implementación del Expediente Judicial Electrónico en la Administración de Justicia	61
3.1.7. Sistema inteligente de asistencia jurídica basado en inteligencia artificial IURIX Mind	61
3.2. Herramientas utilizadas	62
3.3. Casos de Éxito	62
3.3.1. Optimización del rendimiento en la extracción de información jurídica mediante arquitectura basada en microservicios y message brokers	62
3.3.2. Clasificación de textos legales mediante el modelo BERT-MS	65
3.3.3. Implementación de redes neuronales convolucionales para la clasificación automática multietiqueta de documentos legales	68
3.3.4. Extracción inteligente de información legal con BERT-BiLSTM-CRF	71
3.3.5. Sistema Juco-IS para gestión judicial digital	73
3.3.6. Extracción de Hechos Judiciales mediante Aprendizaje de Patrones semi supervisados e Integración de Ontologías	75
3.3.7. Extracción de Entidades en Documentos de Sentencias Judiciales Mediante el Método de Reglas de Expresiones Regulares	77
3.3.8. "Desarrollo del Software 'Radicapp' para la Gestión de Información y Mitigación de la Congestión en los Despachos Judiciales de Colombia	79
3.3.9. Sistema web basado en metodología XP para mejorar la gestión de expedientes en el estudio jurídico Gamión Asociados	81
3.3.10. Sistema web para procuraduría pública en Perú	83
Metodología	86
4.1. Diseño de Investigación	86
4.1.1. Tipo de Investigación	86
4.1.2. Nivel de Investigación	86

4.1.3. Enfoque de Investigación	86
4.1.4. Diseño de Investigación	87
4.2. Población y Muestra	88
4.2.1. Población	88
4.2.2. Muestra	88
4.2.3. Unidad de Análisis	89
4.2.4. Diseño Muestral	89
4.3. Matriz de consistencia	91
4.4. Operacionalización de variables	93
4.5. Técnicas e Instrumentos de Recolección de Datos	96
4.5.1. Técnicas	96
4.5.2. Instrumentos	96
4.6. Técnicas para el Procesamiento de la Información	97
Implementación	98
5.1. Ingeniería Inversa (AS - IS)	98
5.1.1. Conceptualización y Requerimientos del sistema anterior	98
5.1.2. Diseño del sistema anterior	98
5.1.3. Funcionalidades del sistema anterior	99
5.2. Alteración (Transformación)	102
5.2.1. Análisis de Problemas y Limitaciones	102
5.2.2. Propuesta de Mejora	103
5.3. Ingeniería Directa (TO - BE)	105
5.3.1. Conceptualización y Requerimientos del nuevo sistema	105
5.3.1.1. Definición de user stories	105
5.3.1.2. Definición del product backlog	108
5.3.1.3. Definición del sprint backlog	109
5.3.1.4. Cronograma	110
5.3.2. Diseño del nuevo sistema	110
5.3.2.1. Modelo de datos	110
5.3.2.2. Arquitectura General	113
5.3.2.3. Prototipado	114
5.3.3. Implementación del nuevo sistema	115
5.3.3.1. Sprint 1: Servicio AI - Procesamiento de Documentos	115
5.3.3.2. Sprint 2: Módulo de Login y Configuraciones	119
5.3.3.3. Sprint 3: Módulo de Consulta de legajos.	120
5.3.3.4. Sprint 4: Módulo de Recepción de legajos	124
5.3.3.5. Sprint 5: Módulo de Reporte de legajos	127
5.3.3.6. Sprint 6: Módulo de Audiencias	128
Análisis de Resultados	130
6.1. Análisis de Resultados para la población 1	130
6.1.1. Análisis de Confiabilidad	130

6.1.1.1. Dimensión 1: Recepción y Consulta de Expedientes Judiciales	130
6.1.1.2. Dimensión 2: Elaboración y Envío de Recursos Legales	131
6.1.1.3. Dimensión 3: Planificación y Participación en Audiencias	131
6.1.1.4. Dimensión 4: Revisión y Seguimiento de Expedientes	132
6.1.2. Análisis Estadístico Descriptivo	132
6.1.2.1. Análisis por ítems	132
6.1.2.2. Análisis por dimensiones	143
6.1.3. Análisis Estadístico Inferencial	149
6.1.3.1. Análisis de Normalidad	149
6.1.3.2. Contrastación de hipótesis	149
6.1.3.3. Análisis del tamaño del efecto	153
6.2. Análisis de Resultados para la población 2	154
6.2.1. Análisis Estadístico Descriptivo	154
6.2.2. Análisis Estadístico Inferencial	157
6.2.2.1. Prueba de normalidad	157
6.2.2.2. Prueba de U de Mann-Whitney	158
6.3. Resumen de resultados obtenidos	160
6.4. Discusión de Resultados	162
Conclusiones	164
Recomendaciones	165
Limitaciones	167
Referencias Bibliográficas	168
Anexos	176
Anexo 1: Delitos gestionados por la PSI - MININTER	176
Anexo 2: Cuestionario para la población 1	177
Anexo 3: Fichas de registro de datos para la población 2	179
Anexo 4: Validación de expertos	181
Anexo 5: Instrumento 1- Matriz de Datos	182
Anexo 6: Instrumento 1 - Procesamiento de datos	186
Anexo 7: Instrumento 2 - Matriz de datos	188
Anexo 8: Presupuesto del Proyecto	189
Anexo 9: Cronograma del Proyecto	190
Anexo 10: Prompts utilizados	192

Índice de Tablas

Tabla 1: Resultados experimentales sobre la extracción de elementos.	67
Tabla 2: Tabla comparativa de rendimiento de diferentes modelos.	70
Tabla 3: Resultados de Diferentes Modelos al Reconocimiento de Entidades	72
Tabla 4: Rendimiento del enfoque propuesto en el conjunto de datos judiciales	76
Tabla 5: Etapas en el desarrollo del software de gestión de información	80
Tabla 6: Rendimiento del enfoque propuesto en el conjunto de datos judiciales	82
Tabla 7: Indicadores obtenidos para el sistema de gestión de la Procuraduría de San Martín	84
Tabla 8: Matriz de consistencia	91
Tabla 9: Definición operacional de la variable independiente	93
Tabla 10: Definición operacional de la variable dependiente	93
Tabla 11: Necesidades satisfechas en el sistema anterior	98
Tabla 12: Análisis de necesidades satisfechas en el sistema anterior	102
Tabla 13: Mejoras propuestas para el nuevo sistema	104
Tabla 14: Necesidades y user stories en el nuevo sistema	105
Tabla 15: Sprint backlog del nuevo sistema	109
Tabla 16: Alfa de Cronbach para la Dimensión 1	130
Tabla 17: Alfa de Cronbach para la Dimensión 2	131
Tabla 18: Alfa de Cronbach para la Dimensión 3	131
Tabla 19: Alfa de Cronbach para la Dimensión 4	132
Tabla 20: Medidas clave de la Dimensión 1	143
Tabla 21: Medidas clave de la Dimensión 2	144
Tabla 22: Medidas clave de la Dimensión 3	146
Tabla 23: Medidas clave de la Dimensión 4	147
Tabla 24: Prueba de Shapiro-Wilk	149
Tabla 25: Prueba t para la hipótesis general	150
Tabla 26: Prueba t para la hipótesis específica 1	150
Tabla 27: Prueba de wilcoxon para la hipótesis específica 2	151
Tabla 28: Prueba t para la hipótesis específica 3	152
Tabla 29: Prueba de wilcoxon para la hipótesis específica 4	153
Tabla 30: d de cohen	153
Tabla 31: Medidas descriptivas del indicador I1 en el grupo de control y en el grupo experimental	154
Tabla 32: Medidas descriptivas del indicador I2 en el grupo de control y grupo experimental	155
Tabla 33: Prueba de Kolmogorov-Smirnov para Tiempo de Recepción	157
Tabla 34: Prueba de Kolmogorov-Smirnov para Tiempo de Elaboración	158
Tabla 35: Prueba de U de Mann-Whitney para I1	158
Tabla 36: Prueba de U de Mann-Whitney para I2	159
Tabla 37: Resumen de resultados obtenidos	160

Índice de Figuras

Figura 1: Diagrama de problemas del sistema SIE	13
Figura 2: Modelo general de la reingeniería de software	21
Figura 3: Comparación de precisión de servicios OCR	25
Figura 4: Organigrama del MININTER	35
Figura 5: Funcionamiento interno de la PSI	36
Figura 6: Etapas del Proceso Penal	38
Figura 7: Sistemas informáticos de la procuraduría pública del sector interior	39
Figura 8: Pantalla Principal de EJE	40
Figura 9: Pantalla Principal de SINOE	41
Figura 10: Carga procesal de las procuradurías públicas del MININTER en 2024	42
Figura 11: Carga procesal de la Procuraduría del Sector Interior en 2024	43
Figura 12: Expedientes Judiciales por área jurídica	44
Figura 13: Estado de los expedientes judiciales en la Procuraduría del Sector Interior	45
Figura 14: Arquitectura del sistema de comunicación propuesto	64
Figura 15: Diagrama del Sistema Juco-IS	74
Figura 16: Mapa de Circulación de Entidades	79
Figura 17: Arquitectura del sistema anterior	99
Figura 18: Pantalla Legajos del área penal	100
Figura 19: Pantalla Documentos de Ingreso	100
Figura 20: Pantalla Documentos de Salida	101
Figura 21: Pantalla Reportes	102
Figura 22: User Story SIE-12	107
Figura 23: User Story SSIE-35	107
Figura 24: User Story SIE-7	108
Figura 25: Product Backlog del proyecto.	108
Figura 26: Sprint Backlogs	109
Figura 27: Cronograma del Proyecto.	110
Figura 28: Modelo de Datos – Sección 1.	111
Figura 29: Modelo de Datos – Sección 2.	112
Figura 30: Arquitectura de la solución.	113
Figura 31: Prototipo de la solución.	115
Figura 32: Pipeline de Automatización con Enfoque Human-in-the-loop	115
Figura 33: Etapa de Desarrollo	117
Figura 34: Pantalla Inicio de Sesión.	119
Figura 35: Pantalla Mis legajos.	120
Figura 36: Pantalla legajo - Información General.	121
Figura 37: Pantalla legajo – Partes Procesales.	121
Figura 38: Pantalla legajo – Historial de Audiencias.	122
Figura 39: Pantalla legajo – Resultados Finales.	122
Figura 40: Pantalla legajo – Documentos de Ingreso	123

Figura 41: Pantalla legajo – Documentos de Salida	124
Figura 42: Pantalla Recepción legajos	125
Figura 43: Pantalla Nuevo legajo	125
Figura 44: Pantalla Adicionar legajo	126
Figura 45: Pantalla Reporte General	127
Figura 46: Pantalla Seguimiento	128
Figura 47: Pantalla Audiencias	129
Figura 48: Pantalla Mis Audiencias	129
Figura 49: Gráfico de frecuencia de la Pregunta n°1 por pretest y posttest	132
Figura 50: Gráfico de frecuencia de la Pregunta n°2 por pretest y posttest	133
Figura 51: Gráfico de frecuencia de la Pregunta n°3 por pretest y posttest	134
Figura 52: Gráfico de frecuencia de la Pregunta n°4 por pretest y posttest	134
Figura 53: Gráfico de frecuencia de la Pregunta n°5 por pretest y posttest	135
Figura 54: Gráfico de frecuencia de la Pregunta n°6 por pretest y posttest	136
Figura 55: Gráfico de frecuencia de la Pregunta n°7 por pretest y posttest	136
Figura 56: Gráfico de frecuencia de la Pregunta n°8 por pretest y posttest	137
Figura 57: Gráfico de frecuencia de la Pregunta n°9 por pretest y posttest	138
Figura 58: Gráfico de frecuencia de la Pregunta n°10 por pretest y posttest	138
Figura 59: Gráfico de frecuencia de la Pregunta n°11 por pretest y posttest	139
Figura 60: Gráfico de frecuencia de la Pregunta n°12 por pretest y posttest	140
Figura 61: Gráfico de frecuencia de la Pregunta n°13 por pretest y posttest	140
Figura 62: Gráfico de frecuencia de la Pregunta n°14 por pretest y posttest	141
Figura 63: Gráfico de frecuencia de la Pregunta n°15 por pretest y posttest	142
Figura 64: Gráfico de frecuencia de la Pregunta n°16 por pretest y posttest	142
Figura 65: Histograma de la Dimensión 1	143
Figura 66: Histograma de la Dimensión 2	144
Figura 67: Histograma de la Dimensión 3	145
Figura 68: Histograma de la Dimensión 4	147
Figura 69: Comparación de las dimensiones de Gestión de Expedientes Judiciales	148
Figura 70: Distribución del indicador I1 en el grupo de control y en el grupo experimental	154
Figura 71: Distribución del indicador I2 en el grupo de control y grupo experimental	155
Figura 72: Comparación de promedios para I1 y I2 entre el grupo de control y el grupo experimental	156

Capítulo I

Planteamiento del Problema

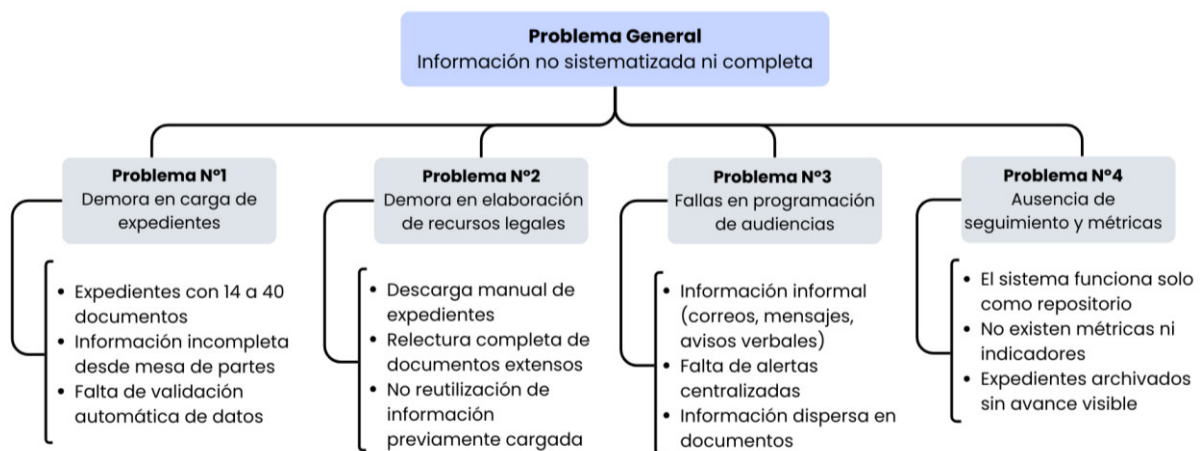
1.1. Descripción del Problema

La gestión de los expedientes judiciales presenta deficiencias asociadas a las limitaciones funcionales y al diseño del sistema web SIE, utilizado por la Procuraduría Pública del Sector Interior para el registro y seguimiento de los procesos judiciales. Si bien el sistema cumple la función básica de repositorio documental, su estructura corresponde a un sistema de tipo legacy, con una organización funcional poco alineada a los procesos operativos reales de la institución. Esta situación genera dificultades en la sistematización de la información, dependencia de actividades manuales, retrasos en las tareas jurídicas y limitada capacidad de seguimiento de los expedientes.

En consecuencia, el sistema no logra apoyar de manera eficiente las actividades relacionadas con la recepción de expedientes, la elaboración de recursos legales, la programación de audiencias y el monitoreo del estado procesal, evidenciando la necesidad de una reingeniería que permita optimizar los procesos y mejorar la gestión integral de la información.

Figura 1

Diagrama de problemas del sistema SIE



Nota. Elaboración Propia

Como se puede apreciar en la Figura 1, los problemas asociados son los siguientes:

- **Problema N.º 1: Demora en la carga de expedientes:** El proceso de recepción y registro de expedientes presenta retrasos debido a la complejidad documental y a la forma en que el sistema organiza la información. Cada expediente está conformado por múltiples documentos, generalmente extensos, cuyo registro inicial es realizado por mesa de partes; sin embargo, la información no siempre se ingresa de manera completa. Esta situación obliga posteriormente a los abogados a revisar manualmente los documentos para obtener información esencial, incrementando el tiempo requerido para iniciar la gestión del expediente, alcanzando un promedio actual de 14 minutos por registro.
- **Problema N.º 2: Demora en la elaboración de recursos legales:** La elaboración de recursos legales se ve afectada por la limitada organización funcional de la información dentro del sistema, así como por la heterogeneidad de los documentos que conforman los expedientes, los cuales suelen presentarse en formatos escaneados, imágenes o anexos fotográficos. Debido a ello, los abogados deben descargar los archivos y realizar una revisión completa para identificar los datos necesarios, sin posibilidad de reutilizar información previamente registrada, lo que incrementa el esfuerzo operativo y reduce la eficiencia del proceso jurídico, registrándose un tiempo promedio actual de 9 minutos por documento elaborado.
- **Problema N.º 3: Fallas en la programación de audiencias:** La programación y seguimiento de audiencias presenta dificultades debido a que el sistema no ofrece una visualización centralizada de la información relacionada con fechas, horarios, enlaces virtuales o ubicaciones físicas, la cual se encuentra dispersa dentro de los documentos del expediente. Como consecuencia, la coordinación depende frecuentemente de comunicaciones informales y de la intervención del personal administrativo, lo que incrementa el riesgo de omisiones, retrasos o inasistencias a diligencias judiciales, afectando la adecuada planificación de las actividades procesales.
- **Problema N.º 4: Ausencia de seguimiento y métricas:** El sistema funciona principalmente como un repositorio documental, sin proporcionar herramientas que permitan analizar el estado de los expedientes o generar indicadores de gestión que faciliten la toma de decisiones. La información no se encuentra estructurada para su explotación, lo que dificulta conocer el avance de los procesos, identificar expedientes sin movimiento o determinar tiempos de

atención, limitando el control oportuno de los casos y afectando la gestión integral de los procesos judiciales dentro de la institución.

1.2. Formulación del Problema

1.2.1. Problema General

¿De qué manera la reingeniería del sistema web SIE permite mejorar la gestión de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER?

1.2.2. Problemas Específicos

- ¿De qué manera la reingeniería del sistema web SIE permite mejorar la recepción y consulta de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER?
- ¿De qué manera la reingeniería del sistema web SIE permite mejorar la elaboración y envío de recursos legales de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER?
- ¿De qué manera la reingeniería del sistema web SIE permite mejorar la planificación y participación en audiencias de los miembros de la Procuraduría Pública del Sector Interior - MININTER?
- ¿De qué manera la reingeniería del sistema web SIE permite mejorar la revisión y seguimiento de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER?

1.3. Objetivos de la Investigación

1.3.1. Objetivo General

Realizar una reingeniería del sistema web SIE para mejorar la gestión de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.

1.3.2. Objetivos Específicos

- Realizar una reingeniería del sistema web SIE para mejorar la recepción y consulta de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.
- Realizar una reingeniería del sistema web SIE para mejorar la elaboración y envío de recursos legales de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.

- Realizar una reingeniería del sistema web SIE para mejorar la planificación y participación en audiencias de los miembros de la Procuraduría Pública del Sector Interior - MININTER.
- Realizar una reingeniería del sistema web SIE para mejorar la revisión y seguimiento de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.

1.4. Delimitación de la Investigación

1.4.1. Delimitación Conceptual

A. Procuraduría Pública del Sector Interior. La Procuraduría Pública del Sector Interior (PSEI) es una entidad pública encargada de la defensa de los intereses jurídicos del Estado en el ámbito de competencia del Ministerio del Interior, conforme a lo establecido en la Constitución Política del Perú y en las normas del Sistema de Defensa Jurídica del Estado. (MINISTERIO DEL INTERIOR, 2019)

En el marco de sus funciones, la PSEI desarrolla actividades de carácter administrativo y judicial, las cuales implican la gestión, registro, seguimiento y control de expedientes judiciales, constituyendo estos procesos un soporte operativo fundamental para el ejercicio de la defensa jurídica del Estado.

B. Sistema web SIE. El Sistema Integral de Expedientes (SIE) se define como el sistema informático de tipo web que utiliza la Procuraduría Pública del Sector Interior para gestionar los expedientes judiciales que son materia de competencia del ministerio, tales como aquellos correspondientes a delitos contra el servicio de seguridad, la integridad institucional y la violación al deber militar policial (V. Díaz, comunicación personal, 12 de marzo de 2023).

No obstante, de acuerdo con el diagnóstico realizado, el sistema presenta limitaciones a nivel de software, principalmente asociadas a un bajo grado de automatización de los procesos documentales, lo que ocasiona que actividades como la recepción, registro, elaboración de documentos, seguimiento y control de expedientes dependan en gran medida de intervenciones manuales, afectando la eficiencia de la gestión judicial.

C. Gestión de expedientes judiciales. La gestión de expedientes judiciales se refiere al conjunto de procesos, métodos y herramientas utilizados para organizar, registrar, tramitar,

consultar, controlar y dar seguimiento a la documentación e información asociada a un proceso judicial dentro de una entidad del sistema de justicia. (Castañeda, 2022)

Esta gestión comprende actividades como la recepción de expedientes, el registro de datos relevantes, la elaboración y envío de documentos procesales, el control de plazos, el seguimiento del estado de los trámites y la consulta o recuperación de información, las cuales deben realizarse de manera oportuna y confiable para garantizar la continuidad del proceso judicial.

En el contexto de la Procuraduría Pública del Sector Interior, la gestión de expedientes judiciales se desarrolla mediante el uso del sistema web interno SIE.

D. Funcionalidades. Conjunto de capacidades o acciones específicas que el sistema puede ejecutar para cumplir con sus objetivos. En el contexto del SIE reingenierizado, las funcionalidades incluyen la carga de documentos, la extracción automática de información, la validación de datos y el almacenamiento estructurado en base de datos.

E. Módulos. Componentes estructurales del sistema que agrupan funciones relacionadas y operan de manera integrada pero independiente dentro de la arquitectura general. Por ejemplo, el módulo de procesamiento OCR y el módulo de análisis semántico mediante LLM.

F. Servicios. Recursos tecnológicos o aplicaciones que proporcionan capacidades específicas al sistema a través de infraestructura externa o interna. Incluyen servicios de almacenamiento, procesamiento, bases de datos o inteligencia artificial que soportan el funcionamiento del SIE.

G. Solución. Conjunto integrado de componentes tecnológicos, procesos y metodologías diseñados para resolver un problema específico. En este caso, la solución corresponde al sistema web SIE reingenierizado, orientado a optimizar la gestión de expedientes judiciales mediante la incorporación de inteligencia artificial.

1.4.2. Delimitación Temporal

La presente investigación se desarrolló en el período comprendido entre abril de 2024 y marzo de 2025, el cual abarca de manera integral las siguientes fases del estudio:

- Fase de investigación y diagnóstico del sistema web SIE, orientada a identificar las limitaciones funcionales y operativas en la gestión de expedientes judiciales. (abril 2024 - agosto 2025)
- Fase de diseño y desarrollo de la reingeniería del sistema de información, de acuerdo con los requerimientos identificados. (septiembre 2024 - diciembre 2024)
- Fase de implementación, pruebas y despliegue. (enero 2025 - febrero 2025)
- Fase de evaluación, que incluye la aplicación del pretest y posttest a los abogados usuarios del sistema y el tiempo de interacción, así como el análisis estadístico de los resultados obtenidos. (marzo 2025)

1.4.3. Delimitación Espacial

Procuraduría Pública del Sector Interior: Jr. Brigadier Mateo Pumacahua N.º 2749, Lince – Provincia Lima– Región Lima –Perú

1.5. Justificación de la Investigación

1.5.1. Justificación teórica

La mejora del desempeño de las instituciones jurídicas constituye una preocupación relevante en diversos países, particularmente en lo relacionado con la gestión de expedientes judiciales y la optimización de los procesos administrativos asociados. En este contexto, los sistemas de información han demostrado ser herramientas clave para facilitar tareas repetitivas, estructurar información y facilitar la toma de decisiones en entornos organizacionales complejos. Diversos estudios internacionales han propuesto soluciones tecnológicas orientadas a la modernización de los sistemas judiciales; por ejemplo, el sistema JuCo-IS desarrollado en Filipinas por Acoba et al. (2020) permitió optimizar los flujos de trabajo de empleados judiciales y abogados en tribunales regionales, mientras que investigaciones como la de Patil y Desai (2023) han planteado la integración de tecnologías como blockchain, inteligencia artificial y aprendizaje automático para reducir cuellos de botella en los procesos judiciales.

En ese sentido, la presente investigación aporta al campo de los sistemas de información aplicados al sector público al analizar cómo la reingeniería del sistema SIE, mediante la incorporación de nuevas herramientas tecnológicas, puede influir en la gestión de expedientes judiciales dentro de una procuraduría pública. El estudio contribuye teóricamente al conocimiento sobre la relación entre

reingeniería de software y desempeño institucional en entornos jurídicos, particularmente en contextos donde predominan documentos no estructurados y procesos administrativos complejos. Asimismo, permite ampliar la evidencia empírica existente sobre el uso de tecnologías emergentes en la optimización de procesos institucionales del sector justicia.

1.5.2. Justificación práctica

Desde el punto de vista práctico, la gestión de expedientes judiciales en la Procuraduría Pública del Sector Interior constituye un proceso crítico para el cumplimiento oportuno de la defensa jurídica del Estado peruano. Sin embargo, el diagnóstico del sistema web SIE evidenció que su funcionalidad se encontraba principalmente orientada al almacenamiento y consulta básica de información, sin incorporar mecanismos que apoyaran o facilitaran procesos manuales como la recepción de documentos, la elaboración de recursos legales, la planificación de audiencias y el seguimiento de expedientes. Esta situación generaba una alta dependencia de actividades manuales, incrementaba los tiempos de procesamiento y dificultaba el control del estado procesal de los casos, especialmente considerando el elevado volumen de expedientes gestionados por la institución.

La reingeniería del sistema SIE permitió incorporar funcionalidades orientadas a facilitar el registro de información, la generación de documentos y la visualización del estado de los expedientes, contribuyendo a mejorar la eficiencia percibida por los usuarios y a reducir los tiempos operativos en determinadas actividades. En consecuencia, los resultados de la investigación pueden servir como referencia para futuras mejoras tecnológicas dentro de la institución y para otras entidades públicas que enfrenten problemáticas similares en la gestión documental y administrativa de procesos judiciales.

1.5.3. Justificación metodológica

Desde el enfoque metodológico, la investigación aporta evidencia empírica mediante la aplicación de un diseño cuasi experimental con mediciones de pretest y posttest en grupos de control y experimental, lo que permitió evaluar el efecto de la reingeniería del sistema sobre la gestión de expedientes judiciales. La utilización de instrumentos de medición validados y el análisis estadístico correspondiente fortalecen la confiabilidad de los resultados y permiten establecer relaciones entre la intervención tecnológica y los cambios observados en las variables de estudio, proporcionando un procedimiento riguroso para analizar mejoras en procesos institucionales.

Asimismo, el estudio incorpora una contribución metodológica al integrar modelos de lenguaje de gran escala (LLM) para la extracción de entidades jurídicas en documentos del ámbito jurídico peruano, junto con un enfoque de recuperación aumentada por generación (RAG) aplicado a documentos legales heterogéneos. Adicionalmente, se propone un esquema de validación humana (human-in-the-loop), en el cual la información procesada por los modelos de inteligencia artificial es posteriormente revisada por el personal de mesa de partes, quien puede validar, corregir o reprocesar los datos extraídos antes de su registro en el sistema.

1.6. Hipótesis

1.6.1. Hipótesis General

La reingeniería del sistema web SIE mejora la gestión de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.

1.6.2. Hipótesis Específicas

- La reingeniería del sistema web SIE mejora la recepción y consulta de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.
- La reingeniería del sistema web SIE mejora la elaboración y envío de recursos legales de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.
- La reingeniería del sistema web SIE mejora la planificación y participación en audiencias de los miembros de la Procuraduría Pública del Sector Interior - MININTER.
- La reingeniería del sistema web SIE mejora la revisión y seguimiento de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.

Capítulo II

Marco Teórico

2.1. Bases Teóricas

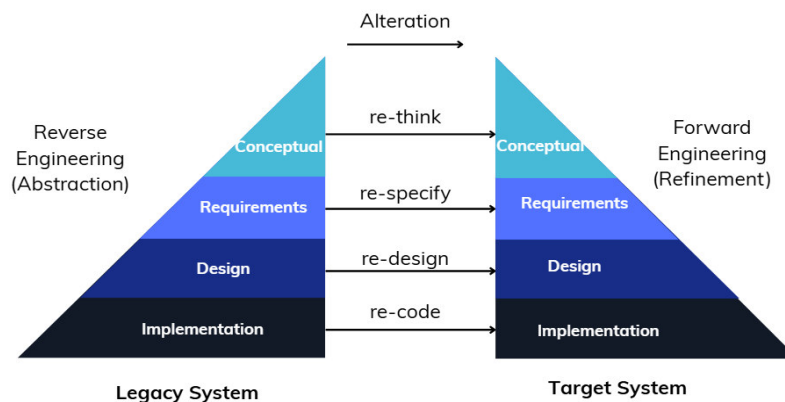
2.1.1. Reingeniería y modernización de software

La reingeniería de software es un proceso orientado a transformar sistemas heredados o legados en sistemas modernizados, con el propósito de adaptarlos a nuevas necesidades tecnológicas, organizacionales y de los usuarios. Este proceso implica el análisis profundo del sistema existente, la comprensión de su estructura y funcionamiento, y la aplicación de mejoras para generar una versión más eficiente, mantenible y alineada con los requerimientos actuales. (Lim & Tan, 2025)

2.1.1.1. Modelo General de la Reingeniería de Software. De acuerdo con Lim & Tan (2025), el modelo general de la reingeniería de software describe el proceso básico mediante el cual un sistema legado es transformado en un sistema renovado. Este modelo está conformado por tres procesos fundamentales: ingeniería inversa, alteración e ingeniería directa, los cuales se desarrollan de manera secuencial e integrada.

Figura 2

Modelo general de la reingeniería de software



Nota. Reproducido de Overview of software re-engineering concepts, models, and approaches, por Lim y Tan (2025), International Journal on Informatics Visualization, 9(1), p. 48.

En la etapa de ingeniería inversa, se analizan los códigos fuente, la arquitectura y la documentación existente con el fin de obtener una representación de alto nivel del sistema.

Este análisis permite comprender su lógica interna, estructura funcional y organización de datos. A medida que avanza esta etapa, el nivel de abstracción del sistema se incrementa, facilitando su interpretación por parte del equipo de desarrollo.

Posteriormente, en la fase de alteración, se realizan las modificaciones necesarias de acuerdo con los nuevos requerimientos. Estas modificaciones pueden incluir la reestructuración del código, la actualización tecnológica, la redefinición de procesos y la reorganización de la arquitectura del sistema. En esta etapa se busca mejorar la calidad interna del software sin afectar negativamente su funcionalidad externa.

Finalmente, la ingeniería directa permite construir el nuevo sistema a partir de la información obtenida en la ingeniería inversa y de las mejoras introducidas durante la alteración. Este proceso inicia con representaciones conceptuales y finaliza con la implementación del software, generando un sistema más eficiente, confiable y mantenible

2.1.2. *Herramientas para la modernización de software*

2.1.2.1. Inteligencia Artificial. La Inteligencia Artificial, también conocida como IA, hace referencia a la creación de algoritmos y plataformas informáticas para realizar tareas que usualmente requerirían la intervención de la inteligencia humana. (Erazo-Luzuriaga, Ramos-Secaira, Galarza-Sanchez y Boné-Andrade, 2023) Estos sistemas están concebidos para razonar, aprender y abordar problemas de forma análoga a la de un ser humano. John McCarthy, uno de los pioneros en este campo, la definió como la ciencia e ingeniería que se dedica a diseñar máquinas con capacidades inteligentes, particularmente programas de cómputo avanzados, que se relacionan con la comprensión de la inteligencia humana natural. (Rodríguez y Barron, 2022)

A diferencia del aprendizaje humano, que se desarrolla de manera gradual a lo largo de años, las máquinas pueden adquirir experiencia de manera rápida, ya que pueden operar las 24 horas del día sin necesidad de descanso ni entretenimiento. La existencia de grandes volúmenes de información provenientes de variadas fuentes y estructuras está estrechamente relacionada con la capacidad de las máquinas para adquirir experiencia al exponerse a una amplia variedad de situaciones y estímulos. (Ávila-Tomás et al., 2020)

2.1.2.2. Reconocimiento Óptico de Caracteres. El proceso de convertir texto manuscrito o mecanografiado en un formato reconocible por computadoras se conoce como Reconocimiento Óptico de Caracteres (OCR). (Khete y Bakshi, 2022) Esta técnica se encuentra ampliamente relacionada con la detección de patrones; asimismo, se concibe como una categorización visual de imágenes digitales, comúnmente identificados con caracteres alfanuméricos. El objetivo es entrenar a la máquina para que reconozca y clasifique estos patrones. (Sarika et al., 2021) En el caso del OCR, el modelo identifica una serie de estructuras definidas por letras, dígitos y otros caracteres especiales. El proceso de entrenamiento implica exponer a la máquina a diferentes tipos de patrones, lo que permite crear un modelo para cada uno de los conjuntos de caracteres. Durante el proceso de reconocimiento, el dispositivo identifica una similitud con los modelos almacenados y asigna el que tenga la mayor similitud. (Esquicha, 2022)

A. Fases de OCR. Según Esquicha (2022), el OCR implica varias etapas clave en su proceso:

Digitalización: Digitalizar es el acto de convertir texto manuscrito o mecanografiado en formato electrónico. (Sarika et al., 2021) Este proceso implica escanear un documento y convertirlo en una imagen electrónica. Para lograrlo, se utilizan escáneres que generan imágenes digitales. Además, se aplica un proceso de umbralización en el que el texto se vuelve negro y el fondo blanco, destacando las partes importantes.

Preprocesamiento: El preprocesamiento mejora la calidad de la imagen escaneada. Esto se logra normalizando la imagen a escala de grises y reduciendo el ruido para obtener una imagen en mapa de bits. (Sarika et al., 2021) Luego, esta imagen se convierte en una imagen diluida, lo que hace que las imágenes sean más legibles para el software. Este proceso es fundamental para imágenes manuscritas, ya que son sensibles al ruido y puede incluir tareas como conversión a escala de grises, binarización, adelgazamiento y normalización. (Mainkar, Katkar, Upade y Pednekar, 2020)

Segmentación: La segmentación implica separar los caracteres reales en una imagen, lo que puede ser complicado, especialmente en documentos manuscritos debido a la

variedad de párrafos, palabras y letras. A menudo, los caracteres adyacentes pueden superponerse o tocarse, lo que dificulta la tarea de segmentación. Este proceso identifica las secciones del documento que evidencian un texto publicado y las distingue de otras formas, como imágenes y gráficos. (Mainkar et al., 2020)

Extracción de Características: Este proceso se identifica como la manera en la que se pueden obtener atributos específicos de los caracteres escritos a mano que permiten su identificación y clasificación. (Sarika et al., 2021) Se utilizan diferentes tipos de características, como diagonales, cruces, puntos finales, características de transferencia, características de zonificación, entre otras. Estas características únicas facilitan el reconocimiento y clasificación de los caracteres durante el procesamiento del texto manuscrito. (Mainkar et al., 2020)

B. Modelos de OCR. En el ámbito del reconocimiento de dígitos escritos a mano, se han investigado diferentes enfoques y modelos. Uno de ellos son los Modelos Tradicionales, que se han enfocado en desarrollar métodos de clasificación para identificar dígitos manuscritos en conjuntos de datos específicos. Estos modelos tradicionales, como k-Nearest Neighbor (KNN), Support Vector Machine (SVM) y otros que no dependen de redes neuronales, han logrado tasas de reconocimiento notables. Por ejemplo, en un estudio se empleó el modelo Local Binary Pattern (LBP) junto con el algoritmo KNN para reconocer dígitos escritos a mano en el contexto de Bengala, logrando una tasa de reconocimiento del 96.7%. (Zhang, 2020)

Por otro lado, han surgido los Modelos de Aprendizaje Profundo, que se destacan en el reconocimiento de dígitos escritos a mano. Estos modelos, como las CNN, han demostrado su eficacia en la identificación de dígitos manuscritos en varios idiomas. Ejemplos de estos modelos incluyen LeNet-5 y CNN. Diversas investigaciones han aplicado estas estructuras para reconocer dígitos escritos a mano en diferentes idiomas, derivando a un aumento sustancial en la precisión del reconocimiento.

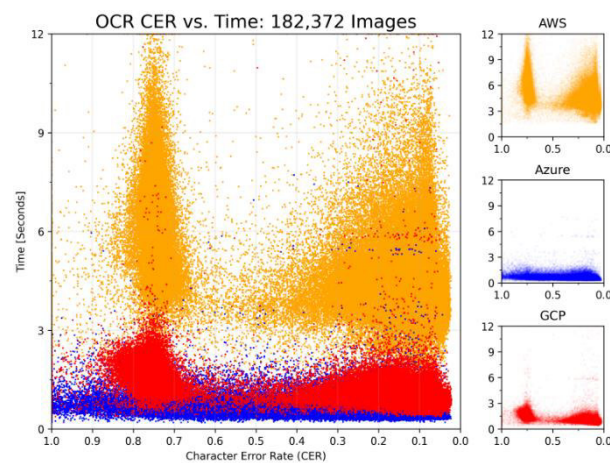
Además de ello, se ha explorado el uso de Modelos de Transferencia de Aprendizaje en el contexto del reconocimiento de dígitos escritos a mano. Estos estudios investigan la posibilidad de mejorar la precisión del reconocimiento al combinar datos de dígitos

manuscritos de diferentes idiomas o conjuntos de datos. Algunos de estos estudios han transferido modelos de entrenamiento entre conjuntos de datos como Urdu, Bangla y Hindi, evaluando si existen características que pueden ser transferidas o aprendidas entre los diferentes conjuntos de datos. Este enfoque ha demostrado ser efectivo para mejorar la precisión del reconocimiento en diversas situaciones.

C. Cloud OCR. Azure Computer Vision, Google Cloud Vision y Amazon Textract son servicios de OCR ofrecidos por las respectivas compañías en la nube. Estos servicios permiten extraer texto y datos de imágenes y documentos, identificar texto en varios idiomas, incluido el manuscrito.

Figura 3

Comparación de precisión de servicios OCR



Nota. Se muestra la comparación de la tasa de error y latencia obtenidas por Azure Computer Vision, Google Cloud Vision y Amazon Textract. Basado en: Ughetta y Kernighan, 2020

En un estudio de comparación, se evaluaron estos servicios utilizando un conjunto de más de 180,000 páginas de registros judiciales históricos, que representaban un desafío para el OCR debido a las variaciones en la ortografía, el diseño y la distinción de impresión a lo largo de más de dos siglos. Además, se compararon con transcripciones humanas para medir su precisión. (Ughetta y Kernighan, 2020)

Los resultados mostraron que Amazon Textract tenía la tasa de error promedio más baja, mientras que Microsoft Azure tenía el tiempo de respuesta promedio más rápido. Google

Cloud Platform (GCP) logró un equilibrio efectivo entre velocidad y precisión. Los resultados también destacaron la similitud entre algunos de los servicios a gran escala, lo que podría deberse a discrepancias entre el contenido de las páginas y las transcripciones humanas.

2.1.2.3. Procesamiento de Lenguaje Natural. El procesamiento de lenguaje natural (NLP) desempeña una función esencial en numerosos sistemas informáticos, desde la revisión de currículos en procesos de contratación hasta la operación de asistentes virtuales y la identificación de contenido no deseado. No obstante, es crucial reconocer que el avance y la aplicación de técnicas de NLP no se distribuye de manera equitativa en todo el mundo, ya que la gran mayoría de los avances en este campo se concentran predominantemente en el idioma inglés. (Abeliuk y Gutierrez, 2021)

Este campo de la inteligencia artificial abarca una amplia gama de aplicaciones y se sustenta en el uso del Machine Learning, Deep Learning y la visión por computadora. Está dedicado a analizar cómo las computadoras procesan y se comunican con el lenguaje natural, lo que habilita la capacidad de procesar, analizar y generar datos en lenguaje humano, abarcando formatos como texto, voz e imágenes. Las aplicaciones de esta tecnología son variadas e incluyen la traducción de idiomas, el análisis de emociones, la clasificación de textos, el reconocimiento de voz y la creación de sistemas de respuesta a preguntas. (Dong, 2023)

A. Etapas de NLP. De acuerdo con Dong (2023), el proceso de procesamiento de lenguaje natural implica una serie de pasos fundamentales que se describen a continuación:

- **Partición de palabras:** La segmentación de palabras en NLP se refiere al proceso de dividir una secuencia continua de texto en unidades lingüísticas significativas, como palabras. Esto es particularmente crucial en idiomas que carecen de separadores de espacio. La segmentación de palabras puede llevarse a cabo mediante diversos enfoques, incluyendo la segmentación basada en diccionario, reglas, modelos estadísticos y aprendizaje profundo entre otros.
- **Etiquetado de partes del discurso:** Este etiquetado consiste en asignar a cada palabra en un texto una etiqueta que indique su función gramatical, como sustantivos, verbos o adjetivos. Esta tarea es fundamental en NLP y es esencial en aplicaciones como la

clasificación de texto, la identificación de elementos y la traducción automatizada. El etiquetado de partes del discurso se realiza comúnmente mediante modelos estadísticos o algoritmos de aprendizaje automático.

- **Análisis gramatical:** El análisis gramatical en NLP tiene como objetivo transformar el texto en lenguaje natural en una estructura más comprensible para las computadoras. Esto implica establecer relaciones gramaticales, como sujeto y predicado. El análisis de sintaxis puede llevarse a cabo mediante enfoques basados en reglas o estadísticas. El enfoque basado en reglas utiliza reglas gramaticales predefinidas, mientras que el enfoque basado en estadísticas aprende patrones empleando grandes volúmenes de datos escritos.
- **Identificación de entidades:** El reconocimiento de entidades es una tarea esencial que se enfoca en identificar y clasificar automáticamente elementos específicos en un texto, como nombres de personas, lugares, organizaciones, fechas y más. Esta tecnología utiliza algoritmos entrenados con datos etiquetados para lograr esta identificación, lo que es fundamental en aplicaciones que requieren extraer información clave de textos.
- **Anotación semántica:** La anotación semántica en NLP implica representar y codificar información semántica en textos en lenguaje natural mediante etiquetas. Esto ayuda a las computadoras a comprender mejor el significado y el contexto de los textos, permitiendo análisis semánticos más precisos.

2.1.2.4. LLMs en el procesamiento de lenguaje natural. Los modelos de lenguaje de gran tamaño (LLM) son modelos probabilísticos entrenados para predecir secuencias de tokens a partir de grandes corpus textuales. Estos modelos pueden generar descripciones, relaciones y estructuras complejas a partir de texto de entrada; sin embargo, también se identifican errores sintácticos y semánticos cuando no se aplican controles adicionales. Frente a modelos tradicionales, los LLM presentan mayor capacidad de representación semántica, pero requieren validación sistemática de sus salidas. (Pontes y Sunyé, 2024)

Desde una perspectiva funcional, los LLM se caracterizan por ser modelos de propósito general, capaces de adaptarse a múltiples tareas lingüísticas sin requerir un

entrenamiento específico para cada una de ellas. Esta versatilidad se debe a que el mismo modelo puede ser aplicado en distintos contextos textuales, ajustando su comportamiento a partir de la información proporcionada en la entrada. (Zhang y Soh, 2024)

A. *LLM aplicados a la extracción de información.* En la extracción de información, los LLM se utilizan como un componente de interpretación semántica que opera sobre texto previamente digitalizado, permitiendo mapear contenido no estructurado a campos estructurados definidos por el sistema. A diferencia de enfoques clásicos basados en expresiones regulares o modelos entrenados para una única entidad, los LLM permiten identificar múltiples tipos de información relevante dentro de un mismo documento, incluso cuando estos aparecen con variaciones de redacción o en distintos contextos. (Ateia et al, 2025) Esta capacidad resulta especialmente útil en documentos judiciales, donde la información clave no sigue un formato uniforme y puede encontrarse distribuida a lo largo del texto.

Adicionalmente, los LLM permiten abordar la extracción de información desde un enfoque orientado a campos definidos por el sistema, sin requerir un entrenamiento supervisado específico para cada tipo de entidad. A partir de descripciones textuales de los campos esperados, el modelo puede identificar y asociar fragmentos relevantes del documento, reduciendo la dependencia de grandes volúmenes de datos previamente anotados. (Adhikary et al, 2024) Este enfoque resulta especialmente relevante en contextos legales, donde los formatos documentales pueden variar entre dependencias o periodos de tiempo, y donde la creación de datasets etiquetados representa un alto costo operativo.

B. *Alucinaciones en modelos de lenguaje.* De acuerdo con el análisis realizado por Kalai et al. (2025), el fenómeno de las alucinaciones en modelos de lenguaje puede definirse y clasificarse a partir de diversos criterios conceptuales y técnicos. En términos generales, las alucinaciones se refieren a errores en los que el modelo genera afirmaciones que resultan aparentemente plausibles, pero que son incorrectas o carecen de sustento factual. En lugar de manifestar incertidumbre o reconocer la ausencia de información suficiente, el modelo produce respuestas con un alto grado de confianza, lo que ha llevado a describir este fenómeno como la generación de “falsedades plausibles con exceso de confianza”.

Desde una perspectiva técnica, los autores sostienen que las alucinaciones no constituyen un comportamiento misterioso o anómalo, sino que responden a la naturaleza estadística del entrenamiento de los modelos de lenguaje. En particular, se argumenta que estas surgen como errores propios de un proceso de “clasificación binaria”, en el cual el modelo no logra distinguir adecuadamente entre afirmaciones verdaderas e incorrectas. Esta limitación se ve acentuada por las presiones estadísticas inherentes al entrenamiento, especialmente cuando la información disponible es ambigua, incompleta o inexistente.

Asimismo, Kalai et al. (2025) proponen una clasificación de las alucinaciones en dos tipos principales:

- **Alucinaciones intrínsecas:** se producen cuando el modelo contradice directamente el prompt o la instrucción proporcionada por el usuario. Un ejemplo de este tipo de alucinación es el conteo incorrecto de letras en una palabra explícitamente indicada en la consulta.
- **Alucinaciones extrínsecas:** ocurren cuando la respuesta generada contradice los datos de entrenamiento del modelo o la realidad externa verificable. Entre los ejemplos más comunes se encuentran la invención de fechas históricas, datos biográficos o títulos académicos inexistentes.

C. Estrategias para la mitigación de alucinaciones. La mitigación de alucinaciones es un desafío crítico para la adopción de modelos de lenguaje (LLM) en entornos profesionales. Las estrategias actuales se dividen fundamentalmente en dos momentos: aquellas aplicadas durante el desarrollo del modelo (preentrenamiento y ajuste fino) y aquellas aplicadas durante la inferencia (cuando el modelo ya está en uso), siendo estas últimas las más accesibles para los usuarios finales (Tonmoy et al., 2024).

Estrategias basadas en el desarrollo y entrenamiento. Una de las formas principales de reducir errores es intervenir en la fase de entrenamiento. El ajuste fino supervisado (Supervised Fine-tuning) ayuda a alinear el modelo con dominios específicos, mientras que el aprendizaje por refuerzo a partir de la retroalimentación humana (RLHF) permite que el modelo aprenda qué tipos de respuestas son veraces y cuáles son inventadas (Tonmoy et al., 2024). Sin embargo, investigaciones recientes sugieren que

el entrenamiento tradicional puede ser insuficiente si no se cambia la forma en que se evalúan los modelos, ya que los sistemas actuales son penalizados por admitir ignorancia, lo que los empuja a "adivinar" e inventar datos para maximizar su puntuación en los exámenes (Kalai et al., 2025).

Estrategias durante la inferencia y el uso de prompts. De acuerdo con Tonmoy et al. (2024), para aplicaciones prácticas, se han desarrollado métodos que no requieren reentrenar el modelo:

- **Generación Aumentada por Recuperación (RAG):** Esta técnica permite al modelo consultar bases de datos externas y documentos de confianza antes de generar una respuesta, lo que reduce drásticamente las alucinaciones al "anclar" el texto a hechos verificables.
- **Cadena de Pensamiento (Chain of Thought):** Consiste en pedir al modelo que explique su razonamiento paso a paso, lo que facilita la detección de errores lógicos antes de llegar a la conclusión final.
- **Auto-corrección y Verificación Multi-agente:** Se emplean múltiples instancias del modelo para que una actúe como generadora y otra como crítica o verificadora, revisando la consistencia de los datos presentados.

Estrategias específicas para analítica de datos y entornos estructurados. En contextos donde se manejan datos precisos (como finanzas o leyes), se han identificado estrategias que superan al ajuste fino tradicional. Estas incluyen la Generación de Salida Estructurada, que obliga al modelo a producir código ejecutable (como SQL) en lugar de texto libre, y la Integración de Capas Semánticas, que traduce el lenguaje natural del usuario a términos estandarizados que el sistema puede procesar sin ambigüedades (Rumiantsau et al., 2024). Además, la implementación de Reglas Estrictas de Aplicación permite configurar al modelo para que se abstenga de responder o pida aclaraciones cuando los datos son insuficientes, mitigando así el riesgo de invención (Rumiantsau et al., 2024).

El factor humano y la responsabilidad profesional. Independientemente de la técnica utilizada, en el ámbito jurídico y profesional persiste el "deber de verificación

humana". Dado que los modelos están optimizados estadísticamente para ser persuasivos incluso cuando mienten, es imperativo que los profesionales certifiquen la exactitud de cada cita y argumento mediante una revisión manual exhaustiva. (Munir, 2025)

2.1.3. Metodología para implementación de LLMs

2.1.3.1. Metodología GenAIOps. GenAIOps (Operaciones de Inteligencia Artificial Generativa) se define como un framework especializado que extiende MLOps para automatizar y gobernar el ciclo de vida completo de modelos de IA generativa como los LLMs, desde la experimentación con prompts hasta el monitoreo continuo de outputs, abordando desafíos únicos como el no determinismo, la evaluación subjetiva y la optimización de recursos en producción. (GeeksforGeeks, 2025)

De acuerdo con GeeksforGeeks, 2025, los pilares fundamentales de GenAIOps incluyen:

- **Gestión de Modelos Fundacionales:** Enfocada en el despliegue eficiente de modelos preentrenados masivos mediante técnicas como fine-tuning (LoRA, QLoRA), prompt engineering y chaining con APIs externas, priorizando adaptación rápida sobre entrenamiento desde cero.
- **Evaluación Generativa Específica:** A diferencia de métricas tradicionales de ML predictivo, GenAIOps mide coherencia factual, relevancia semántica, toxicidad, alucinaciones y satisfacción del usuario mediante benchmarks automatizados y feedback humano.
- **Automatización de Operaciones:** Incluye versionado de prompts/datos/modelos, CI/CD para pipelines generativos, y herramientas como LangFuse para trazabilidad completa de la cadena de generación.
- **Gobernanza y Seguridad:** Implementa guardrails (Guardrails AI), detección de drift semántico y políticas de acceso para mitigar riesgos éticos, asegurando despliegues responsables y escalables en entornos empresariales.

2.1.3.2. Etapas de la metodología GenAIOps. Según Ueno et al. (2024), las etapas de la metodología de GenAIOps son las siguientes:

- ***Planificación.*** La primera etapa corresponde a la fase de planificación, en la cual se define el propósito del proyecto y se decide desarrollar una aplicación que incorpore inteligencia artificial generativa. En este momento se establecen los objetivos, se identifica el problema a resolver y se determina el tipo de tecnología a emplear, como por ejemplo el uso de modelos de lenguaje o sistemas de recuperación de información. Esta etapa es fundamental, ya que orienta todo el proceso posterior.
- ***Desarrollo de la aplicación.*** En esta fase se procede con el diseño y programación de la aplicación. Incluye la selección del modelo de lenguaje más adecuado y la integración de sus funcionalidades dentro del sistema. El equipo desarrolla el código necesario, configura las interfaces y adapta la arquitectura para garantizar un funcionamiento eficiente. Además, se establecen las primeras versiones de los prompts que permitirán interactuar con el modelo.
- ***Pruebas y ajuste de prompts.*** Una vez desarrollada la aplicación, se inicia la etapa de pruebas. En este momento se evalúa el desempeño del sistema y se realizan ajustes mediante técnicas de ingeniería de prompts. El objetivo es lograr que el modelo produzca respuestas coherentes, útiles y acordes con los requerimientos establecidos. Este proceso puede repetirse varias veces hasta alcanzar resultados aceptables.
- ***Publicación o despliegue.*** Después de completar las pruebas, la aplicación se libera para su uso por parte de los usuarios finales. En esta fase, el sistema ya cuenta con prompts optimizados y configuraciones estables. El despliegue marca el inicio de la interacción real con los usuarios, lo que permite observar el comportamiento del sistema en un entorno operativo.
- ***Observación y retroalimentación.*** Una vez en funcionamiento, la aplicación entra en una etapa de monitoreo. Se recopila información sobre el uso, el rendimiento y la satisfacción de los usuarios. Esta retroalimentación permite identificar errores, limitaciones o nuevas oportunidades de mejora. La observación es clave para garantizar la evolución constante del sistema.

- ***Mantenimiento y planificación de nuevas versiones.*** Con base en los datos obtenidos durante la observación, se inicia el proceso de mantenimiento. En esta etapa se analizan posibles actualizaciones, como el cambio de modelo, la incorporación de nuevas funciones o la mejora del rendimiento. También se planifica el desarrollo de versiones futuras que respondan a las nuevas necesidades detectadas.
- ***Adaptación y validación de modelos.*** Cuando se decide cambiar o actualizar el modelo de lenguaje, se realizan pruebas específicas para verificar su compatibilidad con la aplicación. Esto incluye pruebas de funcionamiento básico y validaciones para asegurar que los resultados sean similares o superiores a los obtenidos con el modelo anterior. En muchos casos, es necesario volver a ajustar los prompts.
- ***Pruebas de regresión.*** Las pruebas de regresión permiten comprobar que los cambios realizados no afectan negativamente al comportamiento del sistema. Se comparan los resultados actuales con versiones anteriores para verificar que la aplicación mantiene su calidad. Dado que las respuestas de los modelos pueden variar, estas pruebas se adaptan al contexto de la inteligencia artificial generativa.
- ***Automatización del ajuste de prompts.*** Como parte de la optimización, GenAIOps incorpora técnicas de automatización para mejorar los prompts. Estas herramientas permiten generar o ajustar instrucciones de manera automática, reduciendo la dependencia del trabajo manual. Aunque estas técnicas aún presentan limitaciones, representan un avance importante para agilizar el desarrollo.
- ***Evaluación continua.*** La metodología también contempla una evaluación permanente basada en criterios como funcionalidad, seguridad, confiabilidad y equidad. Se emplean métricas específicas para medir la calidad de las respuestas y detectar posibles sesgos o riesgos. Esta evaluación garantiza que la aplicación cumpla con estándares éticos y técnicos.
- ***Mejora y optimización.*** Finalmente, la información obtenida en todas las etapas anteriores se utiliza para implementar mejoras continuas. Esto incluye la combinación adecuada de técnicas como el prompt tuning, el aprendizaje con pocos ejemplos y la

ingeniería automática de prompts. La optimización busca mantener la competitividad y adaptabilidad del sistema a lo largo del tiempo.

2.1.4. *Gestión de expedientes judiciales en la procuraduría pública del Sector Interior - MININTER*

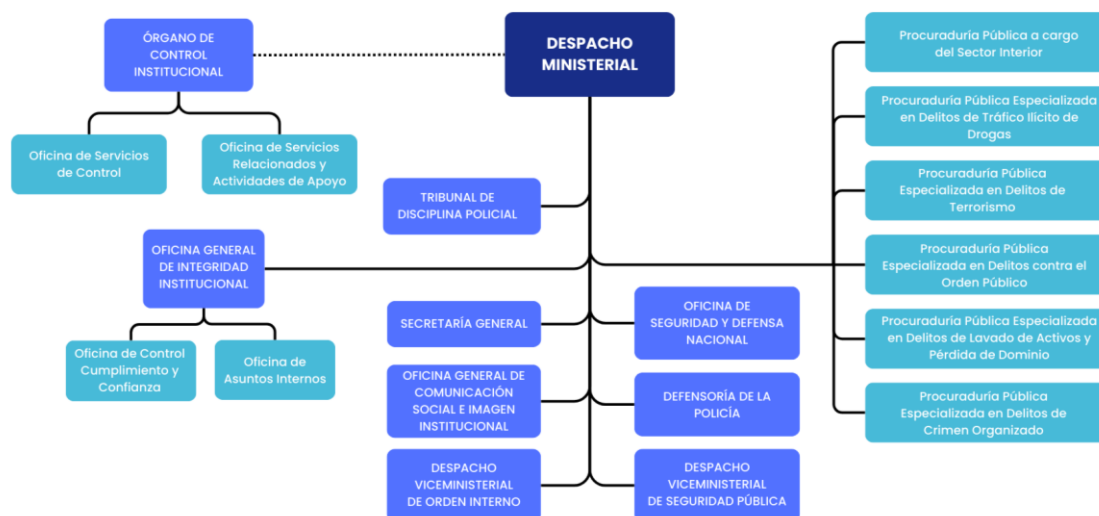
2.1.4.1. Procuradurías públicas del Ministerio del Interior. Las procuradurías públicas conforman el órgano encargado de la defensa jurídica del MININTER. Actualmente, se clasifican en dos grandes grupos. Por un lado, se encuentra la Procuraduría Pública a cargo del Sector Interior, responsable de representar y defender los derechos e intereses del sector. Por otro lado, se encuentran las Procuradurías Públicas Especializadas, las cuales, como su nombre lo indica, intervienen en las investigaciones preliminares, preparatorias y en los distintos procesos vinculados a los delitos en los que están especializadas. (MINISTERIO DEL INTERIOR, 2019)

En la actualidad, el Ministerio del Interior cuenta con cinco procuradurías especializadas:

- Procuraduría Pública Especializada en Delitos de Tráfico Ilícito de Drogas
- Procuraduría Pública Especializada en Delitos contra el Orden Público;
- Procuraduría Pública Especializada en Delitos de Terrorismo;
- Procuraduría Pública Especializada en Delitos de Lavado de Activos y Pérdida de Dominio
- Procuraduría Pública Especializada en Delitos de Crimen Organizado.

Figura 4

Organigrama del MININTER



Nota. Se muestra el organigrama del Ministerio del Interior. Basado en: MINISTERIO DEL INTERIOR, 2019

2.1.4.2. Procuraduría Pública del Sector Interior y sus funciones. Entre las principales funciones de la PSI se encuentran:

- Representar y defender los derechos e intereses del Estado ante los organismos jurisdiccionales, de acuerdo con las normas del Sistema de Defensa Judicial
- Ejercer la plena representación del Estado en los procesos en los que actúe como demandante, demandado, denunciante o parte civil
- Asumir la defensa judicial del Estado en el ámbito del Ministerio del Interior ante cualquier tribunal o juzgado del país.

La Procuraduría del Sector Interior (PSI) se estructura en tres roles principales: el procurador público, los abogados y la secretaría o mesa de partes. El procurador dirige y supervisa los procesos judiciales, presenta denuncias y demandas, informa a las autoridades correspondientes y puede delegar su representación a los abogados auxiliares. (MINISTERIO DEL INTERIOR, 2000). Por su parte, los abogados representan al Estado en audiencias e instancias judiciales, organizándose en áreas especializadas como civil, penal, constitucional,

laboral y contencioso administrativo, siendo el área penal la de mayor carga procesal. (MINISTERIO DEL INTERIOR, 2025)

La secretaría o mesa de partes cumple funciones administrativas esenciales, como la recepción, clasificación, distribución y archivo de documentos. Además, apoya en la redacción de documentos administrativos y en la traducción de documentos provenientes del extranjero, contribuyendo al adecuado funcionamiento de la procuraduría. (MINISTERIO DEL INTERIOR, 2000)

2.1.4.3. Funcionamiento interno de la Procuraduría del Sector Interior.

Figura 5

Funcionamiento interno de la PSI



Nota. Se muestra un diagrama alto nivel del funcionamiento de la PSI. Basado en: MINISTERIO DEL INTERIOR, 2019

El funcionamiento interno de la Procuraduría del Sector Interior se basa en la interacción entre sus áreas internas, las entidades externas y los sistemas de comunicación

utilizados en los procesos penales. La institución recibe documentación proveniente de fiscalías, juzgados y otras entidades del Ministerio del Interior, la cual es gestionada a lo largo de todas las etapas del proceso mediante el uso del legajo como unidad integral del caso.

La recepción, registro y organización de los documentos está a cargo de la mesa de partes, que utiliza el Sistema de Información de Expedientes (SIE) como plataforma central. A partir de esta información, los abogados realizan las actividades de defensa y elaboran los documentos legales correspondientes, los cuales son enviados a las entidades competentes a través del SINOE o del correo institucional, cerrando el ciclo de gestión documental.

2.1.4.4. Rol de la procuraduría en los procesos penales. Las procuradurías públicas cumplen un rol esencial en los procesos penales al representar y defender los intereses del Estado, especialmente en casos que generan perjuicio económico o institucional. Su función principal es procurar la reparación del daño causado y proteger el patrimonio estatal, participando activamente a través de los procuradores públicos. (Río, 2010)

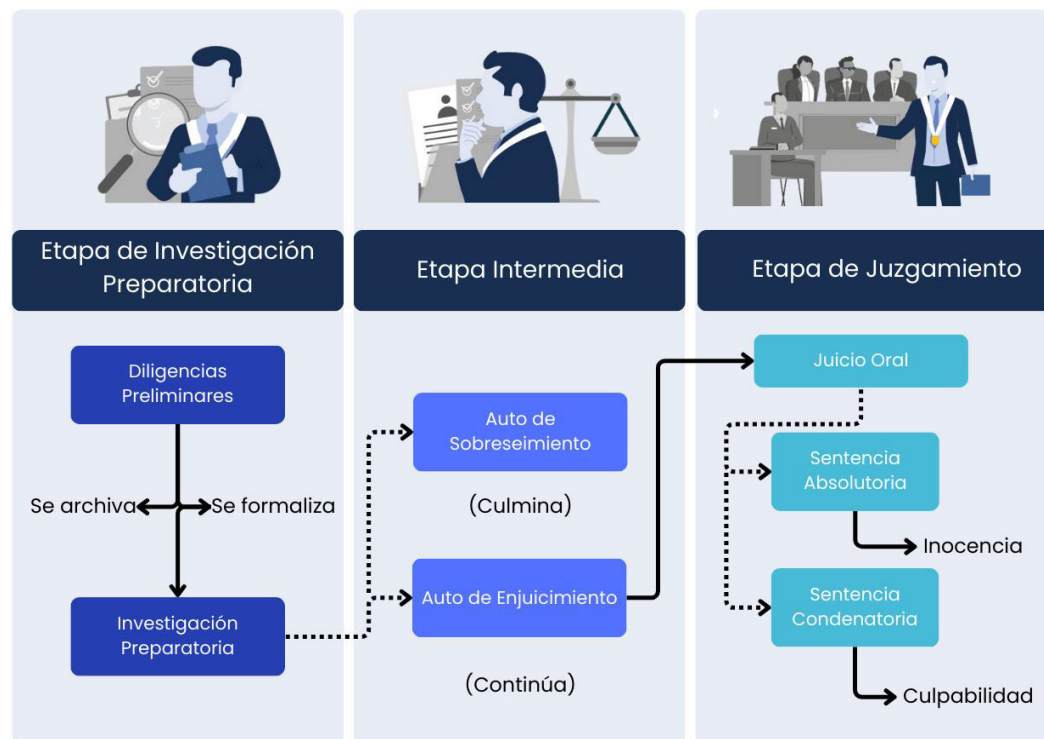
En el proceso penal, las procuradurías intervienen como actor civil, lo que les permite ejercer la acción civil derivada del delito para obtener una reparación por los daños ocasionados. En esta condición, cuentan con diversas facultades procesales, como ofrecer pruebas, participar en las investigaciones, intervenir en el juicio oral e interponer recursos impugnatorios. Asimismo, pueden incorporarse desde las etapas iniciales del proceso, coordinando con el Ministerio Público y fortaleciendo la investigación. (Congreso de la República del Perú, 2004)

Durante la etapa intermedia, la procuraduría presenta su pretensión civil y sustenta el monto solicitado como reparación, además de formular observaciones a la acusación fiscal cuando corresponda. De esta manera, contribuye a que los daños al Estado sean debidamente valorados antes del inicio del juicio oral. (Río, 2010)

2.1.4.5.Etapas principales del proceso penal. El proceso penal se desarrolla a través de diversas etapas orientadas a investigar, juzgar y resolver los hechos delictivos, garantizando el respeto de los derechos fundamentales de las partes.

Figura 6

Etapas del Proceso Penal



Nota. Se muestra el flujograma de decisiones durante el proceso penal. Basado en: MINISTERIO PÚBLICO, 2024

La primera fase es la investigación preparatoria, dirigida por el Ministerio Público con apoyo de la Policía Nacional, y comprende las diligencias preliminares y la investigación formalizada, en las cuales se reúnen los elementos de convicción necesarios y se aplican, de ser necesario, medidas cautelares bajo control judicial. (Oré & Loza, 2005; Ministerio de Justicia, 2020)

La etapa intermedia está a cargo del Juez de la Investigación Preparatoria y tiene como finalidad evaluar los resultados de la investigación. En esta fase, el fiscal decide si formula acusación o solicita el sobreseimiento, y el juez realiza el control correspondiente mediante

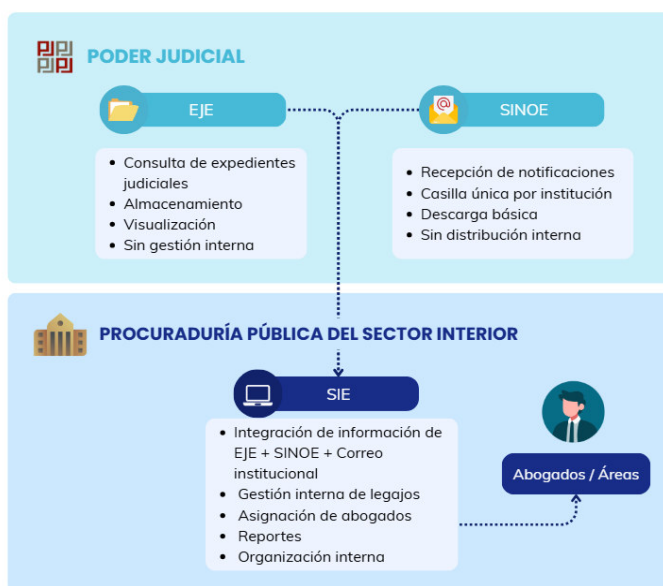
una audiencia, pudiendo disponer el pase a juzgamiento a través del auto de enjuiciamiento. (Oré & Loza, 2005; Ministerio de Justicia, 2020)

Finalmente, la etapa de juzgamiento o juicio oral constituye la fase central del proceso penal y se encuentra a cargo del juez penal. En esta etapa se desarrollan los períodos inicial, probatorio y decisorio, en los cuales las partes presentan sus alegatos, actúan las pruebas y se emite la sentencia condenatoria o absolutoria, según corresponda. (Ministerio de Justicia, 2020; Ministerio Público, 2024)

2.1.4.6. Sistemas informáticos utilizados por la Procuraduría Pública del Sector Interior. La Procuraduría Pública del Sector Interior emplea diversos sistemas externos desarrollados por el Poder Judicial para la gestión de los procesos judiciales, principalmente el Expediente Judicial Electrónico (EJE) y el Sistema de Notificaciones Electrónicas (SINOE).

Figura 7

Sistemas informáticos de la procuraduría pública del sector interior



Nota. Elaboración propia.

a. Expediente Judicial Electrónico (EJE). es un sistema que permite el registro, almacenamiento y consulta de los documentos generados durante un proceso judicial, integrando herramientas como la Mesa de Partes Electrónica, el visor de expedientes, la firma

digital y el sistema de notificaciones. Su principal función es facilitar el acceso a resoluciones, escritos y actuaciones procesales mediante búsquedas basadas en el código del expediente. No obstante, su orientación se centra en la visualización de información, presentando limitaciones en la búsqueda avanzada y en la gestión interna de los casos, por lo que funciona esencialmente como un repositorio digital. (OECD, 2024)

Figura 8

Pantalla Principal de EJE

Nota. Se muestra la pantalla de búsqueda de expedientes de EJE. Basado en: OECD, 2024

b. Sistema de Notificaciones Electrónicas (SINOE). Es una plataforma destinada a la recepción y envío de comunicaciones procesales entre el Poder Judicial y las instituciones públicas, funcionando como una casilla electrónica oficial para la notificación de resoluciones y documentos judiciales (Poder Judicial, 2022). Cuenta con módulos de registro, autenticación y búsqueda básica por expediente y fechas; sin embargo, su diseño contempla una única casilla institucional, lo que centraliza las notificaciones y dificulta su distribución interna, el seguimiento individual de los casos y el control eficiente de los plazos procesales, limitando su utilidad operativa en entidades con alta carga procesal.

Figura 9

Pantalla Principal de SINOE

Búsqueda de Notificaciones

N° Expediente:

Fecha Inicial:

02/04/2022

Fecha Final:

08/04/2022

Estado de Revisión:

Todos

El periodo máximo de búsqueda de las notificaciones es de 30 días.

Buscar

Mover a

Crear Carpeta

Exportar Excel

CARPETAS

CASILLA

ARCHIVADOS

PERSONALIZADO

PRUEBA

SENTENCIAS

Leyenda

No Leído

Leído

Registros: 6 - [Página : 1/1]

1

15

		N°	N° Notificación	N° Expediente	Sumilla	O.J.	Fecha	
		1	35009-2022	05064-2017-0-0901-JR-PE-00	FICHA DE REQUISITORIA + RESOLUCIONES + OFICIO	13° JUZG° UNIPERSONAL (FUNC. LIQUIDADORA) SEDE CENTRAL	06/04/2022 15:44:21	
		2	35008-2022	05064-2017-0-0901-JR-PE-00	FICHA DE REQUISITORIA + RESOLUCIONES + OFICIO	13° JUZG° UNIPERSONAL (FUNC. LIQUIDADORA) SEDE CENTRAL	06/04/2022 15:34:15	
		3	00176-2022	00041-2022-0-3301-JR-PE-02	FICHA DE REQUISITORIA + RESOLUCIONES + OFICIO	2° JUZGADO DE INVESTIGACIÓN PREPARATORIA- VENTANILLA	05/04/2022 16:06:42	
		4	00175-2022	00041-2022-0-3301-JR-PE-02	FICHA DE REQUISITORIA + RESOLUCIONES + OFICIO	2° JUZGADO DE INVESTIGACIÓN PREPARATORIA- VENTANILLA	05/04/2022 15:48:48	
		5	00174-2022	00044-2022-0-3301-JR-PE-01	FICHA DE REQUISITORIA + RESOLUCIONES + OFICIO	2° JUZGADO DE INVESTIGACIÓN PREPARATORIA- VENTANILLA	05/04/2022 14:54:10	
		6	00171-2022	00763-2021-0-3301-JR-PE-02	FICHA DE REQUISITORIA + RESOLUCIONES + OFICIO	2° JUZGADO DE INVESTIGACIÓN PREPARATORIA- VENTANILLA	05/04/2022 12:25:51	

Nota. Se muestra la pantalla de notificaciones de SINOE. Basado en: Poder Judicial, 2022

c. Sistema Integral de Expedientes (SIE). Ante las limitaciones del EJE y del SINOE, se implementó el Sistema Integral de Expedientes (SIE) como una herramienta interna complementaria orientada a fortalecer la gestión documental de la Procuraduría Pública del Sector Interior. Este sistema integra información proveniente de ambos sistemas externos y permite la administración de legajos, la consulta centralizada de expedientes y la generación de reportes, facilitando el seguimiento de los casos y la organización interna (para más información ver la sección 7.1).

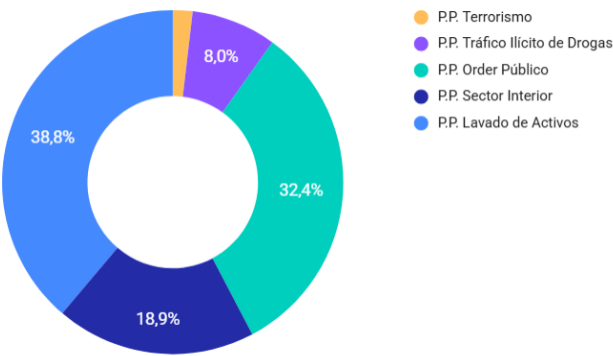
2.1.4.7. Gestión de Expedientes Judiciales en el MININTER.

Según el Ministerio del Interior (2025), la distribución de expedientes judiciales en las procuradurías públicas del sector Interior presenta datos relevantes para el análisis (véase Figura 10). En particular, 3 de las procuradurías manejan una abrumadora cantidad de expedientes judiciales: Procuraduría pública de Lavado de Activos con 47 986 casos (38.8%), Procuraduría pública de Orden Público con 40 032 casos (32.4%) y la Procuraduría Pública del Sector Interior con 23 392 casos (18,9%). Las otras procuradurías, aunque con números

considerablemente más bajos, también presentan un rol significativo en la gestión judicial, centrándose en áreas específicas como delitos de tráfico ilícito de drogas (8.0%) y terrorismo (1.8%).

Figura 10

Carga procesal de las procuradurías públicas del MININTER en 2024

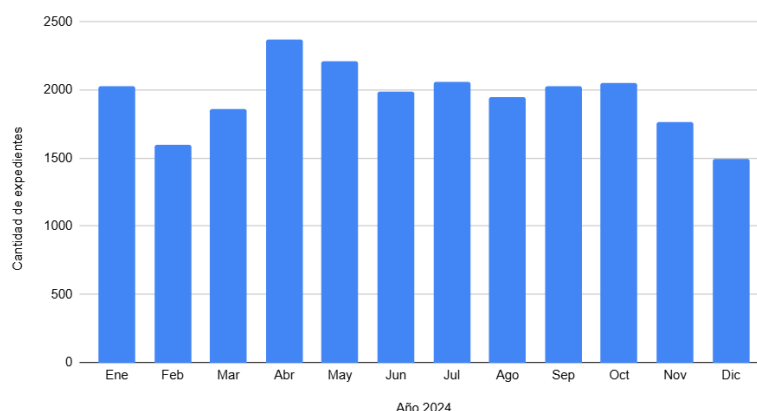


Nota. Se muestra el porcentaje de expedientes judiciales gestionados en 2024. Basado en: MINISTERIO DEL INTERIOR, 2025

Asimismo, la evolución de la cantidad de expedientes judiciales bajo la supervisión de la Procuraduría del Sector Interior durante el año 2024 muestra una tendencia constante que oscila entre 1400 y 2500 expedientes tramitados (véase Figura 11). A lo largo del año, se observa que el mes con mayor cantidad fue abril con 2372 expedientes tramitados y el menor fue diciembre con 1492 expedientes tramitados. Asimismo, a partir de estos datos se obtiene una cantidad promedio 1949 expedientes tramitados y gestionados por mes durante el 2024.

Figura 11

Carga procesal de la Procuraduría del Sector Interior en 2024

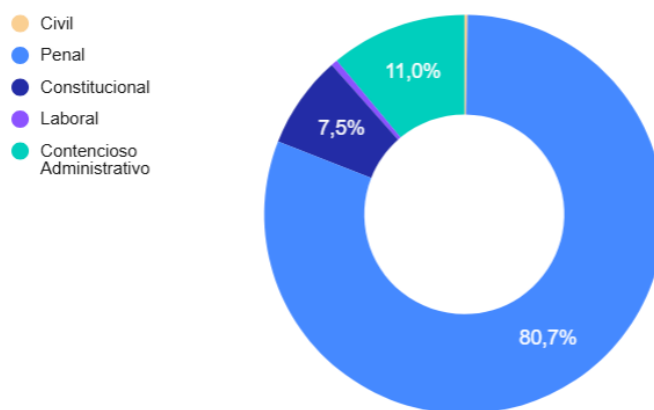


Nota. Se muestra el número de expedientes judiciales gestionados por mes en 2024. Basado en: MINISTERIO DEL INTERIOR, 2025

Por otra parte, durante el 2024, la Procuraduría Pública del Sector Interior registró un total de 23 392 expedientes en diversas áreas legales, incluyendo civil, penal, constitucional, laboral y contencioso administrativo (véase Figura 12). El mayor volumen de casos se concentra en el ámbito penal, con un total de 18 875 expedientes. Esto refleja una alta incidencia de asuntos penales que abordan delitos y procesos legales relacionados. Además, se registran cantidades menores de casos en áreas como el contencioso administrativo, con 2 575 expedientes, constitucional con 18 875 expedientes, laboral con 24 expedientes y civil con 60 expedientes. Esta distribución de expedientes destaca la complejidad y diversidad de casos que maneja la Procuraduría.

Figura 12

Expedientes Judiciales por área jurídica



Nota. Se muestra el porcentaje de expedientes judiciales gestionados por la Procuraduría Pública del Sector Interior según área jurídica en 2024. Basado en: MINISTERIO DEL INTERIOR, 2025

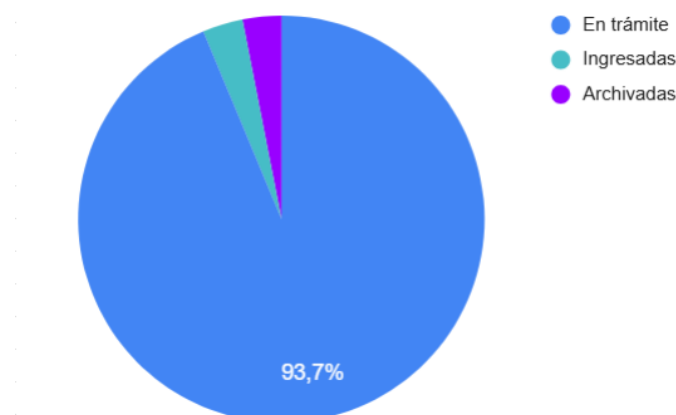
En cuanto al estado de los expedientes (véase Figura 13), se registran 751 expedientes como "Ingresados", lo que indica la cantidad de casos que han ingresado al sistema judicial en 2024 y aún no han sido resueltos. Este número resalta la constante afluencia de nuevos asuntos legales que requieren atención y procesamiento.

Por otro lado, la cifra de "En Trámite" alcanza un total de 21 920 expedientes, lo que demuestra la carga de trabajo en curso que se encuentra en diversas etapas de procesamiento. Esta categoría representa casos que están siendo evaluados, revisados o se encuentran en proceso de resolución. La alta cantidad de expedientes en trámite subraya la necesidad de mantener un flujo de trabajo eficiente y gestionar estos casos de manera efectiva.

Por último, se cuentan 721 expedientes como "Archivados", lo que indica que estos casos se han completado y resuelto de manera satisfactoria. Estos expedientes archivados representan los casos que han pasado por el proceso legal y han alcanzado una conclusión. Este número es esencial para evaluar la eficacia del sistema de justicia en la resolución de casos.

Figura 13

Estado de los expedientes judiciales en la Procuraduría del Sector Interior



Nota. Se muestra el porcentaje de expedientes judiciales gestionados por la Procuraduría Pública del Sector Interior según su estado en 2024. Basado en: MININTER, 2025

2.1.4.8. Informática jurídica. En la actualidad, la incorporación de tecnología en el entorno judicial ha logrado un notable nivel de avance. La digitalización y la automatización han permeado una multitud de actividades relacionadas con el sistema de justicia. Estas aplicaciones informáticas son notables por su diversidad y su repercusión en la calidad y eficiencia del funcionamiento del sistema. Desde la simplificación de la planificación de la agenda de jueces y magistrados, hasta la generación automatizada de documentos legales de alta complejidad, pasando por la gestión de expedientes y la implementación tecnológica identificada como un recurso ideal para respaldar las decisiones vinculadas a fallos judiciales, la tecnología ha revitalizado y transformado el panorama legal. (Sposito, Ledesma, Procopio y Bossero, 2020)

En este contexto, es fundamental destacar la evolución de la llamada "informática jurídica" dentro del Poder Judicial. Esta disciplina se desglosa según Sposito et al., (2020) en tres áreas esenciales que contribuyen a la modernización de las prácticas judiciales:

A. Informática Documentaria. En su esencia, la informática documentaria busca la creación y mantenimiento de un corpus jurídico documentario. Este conjunto de documentos abarca una amplia variedad de fuentes del derecho y tiene como finalidad proporcionar una base sólida para el análisis y la recuperación de información jurídica. Esto permite a los

profesionales legales acceder de manera eficiente a recursos relevantes y fundamentales en sus investigaciones y toma de decisiones.

B. *Informática de Control y Gestión.* En este ámbito, la tecnología se utiliza para supervisar y administrar trámites y procesos judiciales. Esto no solo facilita la actualización constante de la información, sino que también proporciona una perspectiva más esclarecedora y pormenorizada de las actividades judiciales. Como resultado, se logra un mayor control y una gestión más eficiente, lo que a su vez incide en la celeridad y calidad de la administración de justicia.

C. *Informática Decisoria.* La informática decisoria representa una faceta especialmente crucial de la informática jurídica. Se basa en el aprovechamiento de bases de conocimiento jurídico que abarcan diversas áreas del derecho. Sus aplicaciones se extienden a cinco subáreas significativas, entre las que se encuentran el respaldando la elección mediante orientación legal informada, la asistencia en la redacción de documentos legales, la previsión de resultados legales, la investigación jurídica y la educación legal. Estas herramientas tecnológicas permiten a los profesionales del derecho contar con recursos avanzados que mejoran su capacidad para tomar decisiones informadas y para llevar a cabo sus labores de manera más eficiente.

2.1.4.9.E-Justicia. La noción de e-justicia ha sido explorada desde diversas perspectivas. Según Loayza (2022), la e-justicia se podría definir como la administración digital enfocada en atender asuntos judiciales, implicando la introducción de nuevas tecnologías innovadoras que integren a su vez la comunicación en la administración judicial. Por otro lado, Coll y Restrepo (2021) proponen una visión más profunda de la e-justicia, considerándola una nueva manera de estructurar la gestión administrativa en el ámbito judicial proporcionando un enfoque innovador. Esta perspectiva se centra en el conocimiento de las transformaciones sociales contemporáneas y futuras, donde las herramientas de información y de comunicación traen consigo efectos sustanciales en las interacciones humanas, que van más allá de las meras subjetividades y se entrelazan con factores objetivos.

En un enfoque más amplio, Gonzales (2017) citado en Loayza (2022) amplía la definición de e-justicia al abarcar tres áreas distintas. Primero, se centra en la gestión de documentación y la interacción, que abarca la introducción del registro digital. En segundo

lugar, se enfoca en la tramitación de procesos contenciosos, con un énfasis especial en la gestión de pruebas. Finalmente, aborda la conceptualización de documentación electrónica desde el punto de vista procedimental abarca un sinnúmero de audiencias verbales que representan una influencia considerable en la estructura judicial. No obstante, también destaca que la noción conceptual de la judicatura digital es más que la simple aplicación de tecnología. Se respalda en los fundamentos de la ética que evocan un Estado de Derecho, en las que se menciona la disponibilidad de justicia social, la ejecución de decisiones razonable por parte del tribunal, el derecho a presentar pruebas, la declaración de inocencia a menos que haya una prueba evidente que no lo demuestre y el derecho a la participación de un juicio seguro y equitativo.

Por otra parte, Cordella y Contini (2020) plantean que la e-justicia puede entenderse a través de tres sistemas distintos. El primer sistema se relaciona con la gestión de trámites legales y plataformas de respaldo judicial. El segundo abarca el registro digital y la intercomunicación de archivos legales a través de medios electrónicos. Por último, se enfoca en los medios de compatibilidad de la jurisdicción punitiva. Cada uno de estos sistemas tiene consecuencias con distintos efectos para los usuarios, estos se evidencian de forma variada durante el desarrollo de sus labores jurídicas.

2.1.4.10. Aplicación de la IA en el ámbito judicial. La incorporación de la Inteligencia Artificial (IA) al entorno jurídico ha transformado radicalmente la práctica legal. Según Rodríguez et al. (2022), este impacto se manifiesta principalmente en cuatro áreas: la creación de modelos de apoyo para la argumentación y toma de decisiones; la clasificación eficiente de documentos; la extracción ágil de datos en textos complejos; y la planificación de sistemas legislativos basados en tendencias de datos. Estas capacidades convierten a la IA en un recurso estratégico para enfrentar los desafíos legales contemporáneos.

En el plano internacional, un referente destacado es el Sistema Inteligente de Asistencia Judicial (SIAJ) de Shenzhen. Este sistema, integrado en la plataforma digital de los tribunales, emplea modelos de lenguaje (LLMs) entrenados con más de 3 millones de documentos legales para realizar búsquedas semánticas y generar borradores de resoluciones. Liu y Li (2025) explican que el SIAJ opera bajo un modelo de interacción *human-in-the-loop*: la IA actúa como soporte cognitivo proponiendo contenidos preliminares, pero el juez

mantiene la potestad de validar y corregir la decisión final. Esta implementación ha incrementado la productividad judicial entre un 15 % y un 30 %, reduciendo sustancialmente los tiempos de redacción y análisis.

En el contexto latinoamericano, el sistema Prometea, implementado en 2017 por el Ministerio Público Fiscal de la Ciudad de Buenos Aires, ha sido pionero en la automatización de dictámenes. Diseñado para optimizar el "Plan Estratégico de la Fiscalía", su objetivo es resolver la ineficiencia derivada de tareas repetitivas. Estévez et al. (2020) señalan que la gran mayoría de las actividades de la Fiscalía son recurrentes, lo que justificó el desarrollo de modelos predictivos para estandarizar soluciones. Gracias a esto, Prometea ha logrado agilizar significativamente la gestión de casos mediante la comparación de fallos previos y la automatización de decisiones.

Finalmente, derivado de la experiencia de Prometea, surge en Colombia la herramienta PretorIA. Su función principal es apoyar a la Corte Constitucional en la selección y revisión de tutelas, un organismo que recibe más de 2.500 sentencias diarias. La eficiencia de PretorIA es notable: mientras que una persona tardaría aproximadamente 27 minutos en verificar un sumario, el sistema puede procesar miles en segundos. Además de su velocidad, los autores destacan su precisión, alcanzando una efectividad superior al 95 % en los criterios establecidos por la Corte.

2.2. Antecedentes

2.2.1. Antecedentes Internacionales

Ruiz et al. (2020), en España, en su artículo científico sobre optimización de sistemas de gestión documental judicial, propusieron una arquitectura basada en servicios web y brokers de mensajería con el objetivo de mejorar el rendimiento de un sistema de Gestión Electrónica de Documentos (EDMS) que incorporaba módulos de procesamiento de lenguaje natural (NLP). La propuesta fue evaluada en un sistema real de la empresa InSynergy, el cual realizaba análisis lingüísticos de documentos judiciales y presentaba problemas de latencia en escenarios de alta concurrencia. Los autores plantearon desacoplar el subsistema de análisis semántico (AIS) mediante una arquitectura de microservicios con mensajería asíncrona, permitiendo la ejecución paralela en múltiples nodos y reduciendo la carga del sistema principal. La arquitectura incluyó nodos del sistema

central, un middleware de mensajería y un sistema de archivos compartido para el intercambio de documentos y resultados. Las pruebas mostraron una mejora del rendimiento del 77%, reduciendo el tiempo de procesamiento de 450 segundos a menos de 100 segundos para 200 documentos, manteniendo un consumo de memoria similar. Los autores concluyeron que la arquitectura basada en microservicios permitió optimizar significativamente el desempeño, facilitar el escalamiento horizontal y mejorar la mantenibilidad de sistemas judiciales con módulos de inteligencia artificial.

Zhong et al. (2020), en China, en su artículo científico sobre Inteligencia Artificial Legal (LegalAI), analizaron la aplicación de diversos modelos de procesamiento de lenguaje natural para la extracción de información en documentos judiciales, con el objetivo de mejorar la eficiencia en la gestión documental del ámbito legal. El estudio evaluó modelos clásicos como TextCNN, DPCNN, LSTM, BiDAF y BERT, incluyendo una versión especializada con parámetros preentrenados en textos legales chinos denominada BERT-MS. Los autores compararon el desempeño de estos modelos mediante métricas de evaluación, evidenciando que BERT-MS y BERT superaron a los demás enfoques, alcanzando puntuaciones F1 superiores al 75%. En particular, BERT-MS obtuvo valores mínimos de F1 de 84.9%, 79.7% y 81.9% en distintas categorías de extracción, demostrando mayor precisión en el reconocimiento de información legal. Los autores concluyeron que los modelos de lenguaje preentrenados en dominios jurídicos permiten mejorar significativamente la extracción e interpretación de documentos judiciales, destacando el potencial de LegalAI para optimizar procesos legales y reducir la carga operativa institucional.

Ming et al. (2020), en China, en su artículo científico sobre inteligencia artificial aplicada al ámbito judicial, propusieron un modelo de clasificación multietiqueta de documentos legales denominado MTLCNN, basado en una red neuronal convolucional preentrenada (TextCNN), con el objetivo de mejorar la digitalización y categorización automática de documentos judiciales. Previamente, realizaron la digitalización mediante reconocimiento óptico de caracteres (OCR) y la segmentación de palabras con la librería Jieba en Python, generando posteriormente vocabularios con TensorFlow VocabularyProcessor para obtener los vectores de entrada del modelo. El MTLCNN estuvo conformado por capas de embedding, convolución, agrupación máxima y una capa sigmoidea para la predicción. Para la evaluación utilizaron 143 documentos legales con más de 600,000 palabras, logrando un puntaje F1 del 98%, lo que evidenció alta precisión en la clasificación. Los autores concluyeron que el modelo permitió optimizar la automatización en la gestión documental judicial,

demonstrando la utilidad del OCR y del aprendizaje profundo para la organización de información legal.

Gu et al. (2020), en Shanghái, China, en su artículo científico sobre extracción de información judicial mediante inteligencia artificial, propusieron un modelo basado en la integración del modelo de lenguaje preentrenado BERT con redes BiLSTM y una capa CRF, con el objetivo de mejorar el reconocimiento de entidades en documentos legales. El sistema estuvo conformado por tres etapas: obtención de la representación semántica mediante BERT, codificación contextual con BiLSTM y generación de etiquetas óptimas mediante CRF. Para la experimentación recopilaron 294 documentos con aproximadamente 160,000 palabras provenientes de diversas fuentes judiciales, además de un diccionario legal de dos millones de palabras basado en legislación china, dividiendo los datos en conjuntos de entrenamiento y prueba con una proporción 75-25. Las entidades fueron clasificadas en cuatro categorías: documento jurídico, sujeto jurídico, objeto y hecho jurídico, comparando el rendimiento con modelos BiLSTM y BiLSTM-CRF tradicionales. Los resultados mostraron una precisión de 88.86%, sensibilidad de 87.49% y F1 de 87.97%, evidenciando que la incorporación de BERT incrementó el rendimiento en 6.42% respecto a modelos sin preentrenamiento. Los autores concluyeron que la combinación BERT-BiLSTM-CRF permitió una mejor representación semántica y mayor precisión en la extracción de entidades judiciales complejas.

Acoba et al. (2020), en Filipinas, en su artículo científico sobre el sistema Juco-IS (Judicial Court – Information System), desarrollaron e implementaron una solución web orientada a optimizar la gestión de expedientes judiciales en un Tribunal Regional de Primera Instancia, mediante la digitalización y automatización de procesos como el registro de casos, la programación de audiencias, la presentación de documentos y el cierre de expedientes. El estudio describió que el sistema estuvo conformado por cuatro subsistemas principales: administración de casos, organización de colas judiciales, registro y transcripción de audiencias, y presentación electrónica de documentos, lo que permitió gestionar integralmente el ciclo de vida del expediente. Asimismo, señalaron que la solución fue construida bajo el patrón arquitectónico Modelo-Vista-Controlador (MVC), facilitando la escalabilidad, mantenibilidad y separación de responsabilidades, además de contemplar distintos perfiles de usuario con niveles diferenciados de acceso a la información. Los autores concluyeron que la implementación del sistema redujo significativamente las tareas manuales y los tiempos administrativos, mejorando la eficiencia operativa institucional.

Anu y Sangeetha (2021), en su artículo científico sobre procesamiento de lenguaje natural aplicado al ámbito judicial, propusieron un modelo basado en aprendizaje semi supervisado mediante la técnica de bootstrapping para la extracción automática de entidades en documentos legales. El modelo fue entrenado inicialmente con un conjunto reducido de datos, identificando patrones asociados a doce categorías predefinidas, los cuales posteriormente se utilizaron como patrones semilla para procesar nuevos documentos. La entrada del sistema consistió en textos judiciales y la salida en archivos JSON estructurados con información de entidades legales. El proceso incluyó etapas de preprocesamiento, normalización del texto, etiquetado de entidades y aprendizaje de patrones para la extracción de hechos candidatos. El experimento utilizó 14 datos semilla y 20 documentos judiciales electrónicos, obteniendo resultados destacados con puntajes F1 de hasta 100% en categorías como número de caso, juez y demandantes, y superiores al 93% en otras categorías. Los autores concluyeron que el enfoque bootstrapping permitió una alta precisión en la extracción de información compleja, demostrando su eficacia incluso con conjuntos de datos reducidos para el entrenamiento.

Kainan y Xin (2021), en China, en su artículo científico sobre procesamiento de lenguaje natural aplicado al ámbito judicial, presentaron un subsistema para el reconocimiento de entidades nombradas (NER) orientado a la extracción automática de información relevante en documentos judiciales, como imputados, lugares, delitos y funcionarios legales. Para el desarrollo utilizaron documentos provenientes del portal Chinese Judgments Online, seleccionados debido a su formato estandarizado, lo que facilitó el procesamiento automático. Los autores emplearon técnicas de expresiones regulares como método principal de extracción, argumentando que la estructura de los documentos judiciales presentaba patrones constantes organizados en secciones definidas, con mínimas variaciones lingüísticas. Las entidades extraídas abarcaron información legal básica, organizaciones judiciales y personal jurídico, cuyos resultados fueron representados mediante mapas de relaciones y una ontología de información del caso para estructurar el conocimiento obtenido. Los autores concluyeron que el uso de expresiones regulares resultó eficiente para la extracción de información en documentos con patrones definidos, demostrando que esta técnica permite reducir la complejidad computacional y optimizar el procesamiento de datos jurídicos estructurados.

Rojas et al. (2022), en Colombia, en su artículo de investigación orientado al ámbito judicial, desarrollaron un sistema web con el propósito de descongestionar los juzgados,

específicamente el Juzgado N.º 1 Civil del Circuito de Fusagasugá, Cundinamarca, buscando reducir el tiempo de atención a los usuarios, mejorar la gestión de expedientes judiciales y automatizar el envío de reportes estadísticos al Sistema Estadístico de la Rama Judicial. El estudio fue de tipo evaluativo con enfoque deductivo preexperimental y se caracterizó por incorporar funcionalidades avanzadas, como la generación automática de documentos legales y el uso de herramientas de visualización de datos para producir indicadores y métricas que permitieron evaluar el desempeño institucional y mejorar el control de actividades diarias. Desde el punto de vista técnico, los autores implementaron una aplicación web de código abierto con una base de datos relacional alojada en servicios de nube, tanto para el servidor de datos como para el servidor web, garantizando disponibilidad, escalabilidad y continuidad operativa ante fallos. Los resultados evidenciaron que la plataforma disminuyó el tiempo de atención al público y redujo los tiempos de respuesta del juzgado frente a los procesos judiciales. Los autores concluyeron que el uso de tecnologías de cloud hosting y dashboards de visualización contribuyó significativamente a optimizar la gestión de información judicial y la toma de decisiones institucionales.

2.2.2. Antecedentes Nacionales

Castañeda (2022), en Perú, en su tesis aplicada desarrollada en la Procuraduría Pública Regional de San Martín, evaluó el impacto de un aplicativo web en la gestión de expedientes judiciales, utilizando la metodología ágil Scrum para el desarrollo del software. La investigación analizó siete indicadores de eficiencia operativa antes y después de la implementación, evidenciando que previamente los procesos se realizaban de forma manual o mediante sistemas no integrados, generando altos tiempos de respuesta en consultas de información externa. Tras la implementación, los tiempos de consulta se redujeron de varias horas a un promedio de 54 segundos, mostrando una mejora significativa en la eficiencia institucional. Asimismo, se reportó un incremento en la proporción de reclamaciones gestionadas del 62.79% al 92.79% y en la tasa de apelaciones del 67.43% al 86.22%, reflejando una optimización en el tratamiento de procesos judiciales. La satisfacción de los usuarios también aumentó, con mejoras en los niveles de aceptación sobre usabilidad y facilidad de uso según escalas Likert. El autor concluyó que la implementación del sistema web contribuyó significativamente a mejorar la gestión de expedientes y el desempeño institucional, destacando la utilidad de Scrum para el desarrollo iterativo en entornos públicos.

Matos (2022), en Perú, en su tesis aplicada desarrollada en el Estudio Jurídico Gamión y Asociados, propuso la implementación de un sistema web para mejorar la administración de expedientes judiciales, a partir del diagnóstico de una elevada carga procesal y discrepancias entre los plazos legales y los tiempos reales de atención. La investigación tuvo un enfoque cuantitativo de tipo experimental, desarrollándose una plataforma con PHP, JavaScript, MySQL y el framework Laravel, empleando la metodología Extreme Programming (XP) y once historias de usuario organizadas en cinco módulos funcionales. Los resultados evidenciaron mejoras significativas en los tiempos operativos, reduciéndose el registro de expedientes de 13.3 a 2.4 minutos, el procesamiento de informes de 20.3 a 1 minuto y la búsqueda de información de 48.6 a 0.91 minutos, representando mejoras superiores al 80%. Asimismo, la percepción de los usuarios calificó el sistema mayoritariamente como “bueno” y “excelente”, asociando la optimización a la automatización de procesos críticos. El autor concluyó que la implementación del sistema web permitió mejorar de manera objetiva la eficiencia en la gestión de expedientes judiciales mediante la reducción de tiempos y la optimización de procesos.

2.3. Marco Conceptual

- **Sprints:** Son ciclos de hasta un mes con inicio y fin definidos. Cada sprint debe producir algo de valor tangible para el cliente. Todos los sprints, por lo general de la misma duración, siguen inmediatamente a su predecesor. Se evitan cambios en objetivos o personal durante un sprint, aunque circunstancias comerciales pueden hacerlo necesario. (Rubin, 2012)
- **Base de datos:** Constituye una agrupación de datos interconectados. Con el término "datos", se abarcan los hechos que son conocidos, susceptibles de ser registrados y que llevan un significado implícito. Por ejemplo, nombres, números de teléfono y direcciones de personas. (Oracle, 2020)
- **Product Backlog:** Se denomina Product Backlog a la lista jerarquizada de trabajos a completar. Contiene características necesarias para cumplir con el desarrollo del producto de software. La lista es dinámica y se ajusta a medida que cambian las condiciones comerciales o el entendimiento del equipo sobre el producto. (Rubin, 2012)
- **Prompt:** Un prompt es la entrada textual que un usuario proporciona a un modelo de inteligencia artificial para guiar la generación de una respuesta. Puede consistir en

instrucciones o preguntas simples, así como en descripciones más complejas según la tarea y el tipo de modelo. En el caso de los LLM, los prompts avanzados pueden estructurarse para orientar el razonamiento del modelo y mejorar la calidad de la respuesta. (Amatriain, 2024)

- **Derecho Civil:** Corresponde a una rama del derecho que regula las conexiones entre ciudadanos y entidades de carácter privado, estableciendo las normas y principios que rigen asuntos de carácter no penal. Este campo del derecho abarca una amplia variedad de temas, incluyendo contratos, propiedad, sucesiones, familia y responsabilidad civil. Su objetivo principal es proporcionar un conjunto normativo para la resolución de controversias y la salvaguarda de los derechos y obligaciones de las partes involucradas en asuntos de naturaleza no criminal. (Rojina, 2008)
- **Derecho Penal:** También conocido como derecho criminal, corresponde a una rama legal enfocada en las infracciones y vulneraciones normativas de las leyes penales y las ramificaciones legales que se desprenden de dichas violaciones. Su principal objetivo es sostener la armonía y el resguardo en la sociedad al castigar a aquellos individuos que han cometido delitos considerados criminales, como homicidio, robo, agresión, fraude y otros. (García, 2023)
- **Estado de Derecho:** Desde un punto de vista riguroso, su definición alude a cualquier sistema en el que las autoridades poseen capacidades estipuladas por una normativa y por ende se encuentran desempeñadas a través de procedimientos legales emprendidos. A su vez, en relación con un aspecto más jurídico, se puede conceptualizar como un método de gobierno sujetado por el derecho que presenta la legalidad de cualquier estado. (Aguiló, 2019)
- **Expediente judicial:** Conjunto ordenado de documentos, actuaciones, escritos y resoluciones que registran el desarrollo de un caso judicial. (RAE, 2025) En el contexto de la gestión documental de la Procuraduría Pública del Sector Interior, se utiliza el término “legajo” para referirse al expediente, tanto en la etapa de diligencias preliminares como cuando el proceso ya ha sido formalizado. Por este motivo, a nivel de sistema, se emplea el término legajo como equivalente de expediente judicial.
- **Hiper Parámetros:** Son valores de configuración externa empleados por los científicos de datos para manejar y refinar el proceso de entrenamiento de modelos de Machine Learning. A

menudo se les conoce como hiper parámetros de modelos, y se configuran manualmente antes de iniciar el entrenamiento de un modelo. (Amazon, 2020)

- **Prompt Engineering:** El prompt engineering es la disciplina orientada al diseño y optimización de entradas textuales para guiar eficazmente el comportamiento y las respuestas de los modelos de IA generativa. Su objetivo es formular prompts que permitan alcanzar resultados específicos, considerando las capacidades, limitaciones y el contexto de uso del modelo. Este enfoque combina conocimiento del dominio, comprensión del modelo y técnicas estructuradas para adaptar los prompts a distintas tareas y escenarios. (Amatriain, 2024)
- **JSON:** JSON, o Esquema de Objetos JavaScript, se encuentra expresado por un formato que permite el intercambio de información a través de un formato ligero. Su legibilidad y facilidad de escritura son notables para las personas, y su capacidad para ser analizado y generado por las máquinas lo convierte en una opción eficaz. Este formato está regido por unas series de estándares emitidos por el lenguaje de JavaScript, específicamente el ECMA-262. JSON. (IBM, 2022)
- **Azure App Service:** Es un servicio en la nube que permite desplegar y ejecutar aplicaciones web, servicios backend y API sin necesidad de gestionar la infraestructura. Proporciona un entorno administrado que simplifica el hospedaje y permite enfocarse en el desarrollo de la aplicación. Además, es compatible con múltiples tecnologías y lenguajes, tanto en entornos Windows y Linux como mediante contenedores personalizados. (Microsoft, 2025)
- **Microservicios:** Representan un enfoque arquitectónico que desglosa una aplicación en una serie de servicios independientes, cada uno con una funcionalidad específica. Esta descomposición permite a los equipos de desarrollo trabajar en módulos separados, lo que conlleva una mayor agilidad y escala en la producción y puesta en operación del software. (Waseem, Liang, Shahin, Di Salle y Márquez, 2021)
- **IA Generativa:** Es un sistema basado en máquinas que opera con distintos niveles de autonomía y que, a partir de los datos de entrada que recibe, es capaz de inferir y generar resultados nuevos como contenidos, predicciones, recomendaciones o decisiones, con impacto en entornos físicos o virtuales. Estos sistemas pueden adaptarse tras su despliegue y persiguen objetivos explícitos o implícitos. (Palomino, Herrera y García, 2025)

- **Patrón Arquitectónico:** Es un diseño probado y ampliamente aceptado que proporciona una solución estructurada y eficiente para problemas comunes en el diseño de sistemas y aplicaciones de software. Estos patrones son fundamentales en el progreso de software ya que ofrecen un enfoque establecido para resolver desafíos recurrentes, como la gestión de datos, la comunicación entre componentes y la organización de la lógica de una aplicación. (Mejía, Agudelo y Gil, 2021)
- **Patrones de semilla:** En el contexto de la inteligencia artificial y el procesamiento de lenguaje natural (NLP, por sus iniciales) se refieren a un conjunto inicial de reglas, etiquetas o expresiones que se utilizan para entrenar o iniciar la detección y extracción de información de texto no estructurado. Estos patrones sirven como base para implementar sistemas de NLP capaces de identificar y categorizar información específica en documentos textuales. (Singh, 2018)
- **Scrum:** Es un enfoque de trabajo ágil que se adopta con frecuencia para la gestión de proyectos, principalmente en software. El método se basa en la colaboración, la apertura y la entrega incremental de valor a través de ciclos denominados Sprints, con una duración típica de 2 a 4 semanas. Scrum promueve la flexibilidad y la adaptación constante, con roles claramente definidos, como el Scrum Master y el Product Owner, y eventos regulares, como la Reunión Diaria de Scrum y la Revisión del Sprint. (Schwaber y Sutherland, 2020)
- **Servicio de Nube:** Es un paradigma innovador que facilita la entrada generalizada y cómoda mediante una red que otorga el establecimiento de información compartida y adaptable. Estos recursos abarcan desde infraestructura, como servidores y almacenamiento, hasta aplicaciones y servicios diversos. Lo distintivo de la nube es su capacidad de aprovisionar y liberar estos recursos de manera ágil y eficiente, sin requerir una gestión complicada por parte del proveedor de servicios ni una interacción extensiva con los usuarios. (Sánchez, 2021)
- **Términos ontológicos:** En el contexto del procesamiento de lenguaje natural (NLP), los términos ontológicos se caracterizan por representar el conocimiento como agrupaciones de conceptos que incluyen sus respectivos atributos, valores y relaciones. Para lograr una definición entendible y generalizada por todos los usuarios, las ontologías requieren tres componentes esenciales: conceptos de dominio, relaciones e instancias. (Chen y Luo, 2019)

- **User Story:** En Scrum, una user story consiste en una práctica utilizada para expresar requisitos funcionales desde la perspectiva del usuario final o del cliente. Cada user story describe un elemento o prestación del producto o sistema considerado importante por el usuario que lo emplea. (Verheyen, 2019)

Capítulo III

Estado del Arte

3.1. Sistemas o Aplicaciones Comerciales

En el ámbito global, el desarrollo de sistemas informáticos orientados a la gestión de expedientes judiciales ha evolucionado constantemente. Estas aplicaciones buscan mejorar la eficiencia operativa mediante la digitalización de procesos, la automatización completa o parcial de tareas administrativas y el control sistemático de la información asociada a los expedientes judiciales. Su implementación responde a la necesidad de reducir tiempos de respuesta, optimizar la organización documental y facilitar el seguimiento de los procesos legales en instituciones con alta carga procesal.

Los sistemas y aplicaciones comerciales existentes incorporan diversas funcionalidades, tales como registro digital de expedientes, gestión documental, control de plazos procesales, programación de audiencias, monitoreo del estado de los casos y generación de reportes estadísticos. En muchos casos, estas soluciones también integran tecnologías emergentes como servicios en la nube, automatización de flujos de trabajo e inteligencia artificial, lo que permite mejorar la trazabilidad de la información y la toma de decisiones institucionales. A continuación, se presentan los principales sistemas identificados en la literatura y en el mercado, resaltando sus características más relevantes y su aporte en la optimización de la gestión de expedientes judiciales.

3.1.1. Sistema de gestión judicial con inteligencia artificial Onesait Government Justice en la Corte de Cajamarca. Minsait (2019). Perú.

La empresa Minsait implementó el sistema Onesait Government Justice en la Corte Judicial de Cajamarca, Perú, como parte de un proyecto piloto orientado a modernizar la gestión de expedientes judiciales mediante inteligencia artificial. El sistema fue desarrollado, desplegado y probado durante dos meses, permitiendo integrar expedientes electrónicos en los equipos de los jueces y sugerir automáticamente párrafos relevantes y referencias legales durante la revisión de casos. La solución utiliza procesamiento de lenguaje natural para organizar información jurídica no estructurada y mejorar la eficiencia en la tramitación documental. Los resultados reportados indican un incremento de hasta 50 % en la velocidad de procesamiento de información, contribuyendo a reducir la carga procesal y el uso de papel en los tribunales. El estudio concluye que este tipo de herramientas

representa un avance significativo en la modernización del sistema judicial peruano y un referente para la automatización de la gestión de expedientes en instituciones públicas.

3.1.2. Sistema de seguimiento y centralización de expedientes judiciales CaseTracking. Lemontech (2021).

El sistema CaseTracking fue desarrollado como una plataforma orientada a centralizar la información procesal y facilitar el seguimiento de expedientes judiciales mediante una interfaz única. La solución ha sido adoptada por diversas entidades públicas en América Latina, incluyendo organismos del Poder Judicial de países como México, Chile, Perú, Colombia y el Consejo de la Judicatura de Ecuador, evidenciando una tendencia regional hacia la digitalización del control procesal. Entre sus funcionalidades principales destacan el acceso en línea a expedientes mediante sincronización automática o carga manual de documentos, la organización de información relevante del caso (estado, órgano jurisdiccional, etapa procesal, número de expediente) y la gestión documental compartida entre usuarios. No obstante, el sistema se centra en la trazabilidad y administración de documentos, sin incorporar mecanismos avanzados de automatización para el análisis del contenido jurídico. El estudio concluye que este tipo de plataformas contribuye a reducir tiempos de búsqueda y mejorar el control administrativo, aunque el procesamiento e interpretación de la información continúe dependiendo del análisis manual de los operadores jurídicos.

3.1.3. Sistema de gestión judicial integral Atlanta Web para la digitalización del proceso jurisdiccional. Gobierno de Canarias (2022). Consejería de Administraciones Públicas, Justicia y Seguridad. Islas Canarias, España.

En el ámbito del Poder Judicial, Atlanta Web fue desarrollado como una solución tecnológica orientada a la digitalización integral del proceso jurisdiccional, desde el ingreso de causas hasta su resolución, promoviendo el expediente judicial electrónico y la eliminación del uso de papel. El sistema incorporó tecnologías como inteligencia artificial, robotización de procesos y aprendizaje automático con el objetivo de optimizar la tramitación judicial y reducir los tiempos de respuesta. Entre sus principales funcionalidades se incluyeron el acceso remoto seguro mediante firma electrónica, un visor del expediente judicial electrónico y herramientas para la consulta y gestión de actuaciones procesales. Asimismo, la plataforma buscó mejorar la relación entre los órganos jurisdiccionales y los ciudadanos mediante servicios digitales integrados. No obstante, estas soluciones estuvieron diseñadas para el Poder Judicial, cuya función es la administración de justicia,

mientras que las procuradurías públicas cumplen un rol de defensa jurídica del Estado; por ello, su transferencia directa es limitada, aunque aporta referentes relevantes en gestión documental, trazabilidad de expedientes y automatización de procesos administrativos vinculados al ciclo de vida de los casos judiciales.

3.1.4. Sistema de gestión de casos judiciales para la administración integral del ciclo de vida procesal. Court Services Victoria (2022). Tribunal de Magistrados de Victoria y Tribunal de Niños de Victoria. Australia.

El *Case Management System* (CMS) desarrollado para el Tribunal de Magistrados de Victoria y el Tribunal de Niños de Victoria constituyó una iniciativa de modernización orientada a reemplazar sistemas judiciales heredados con más de treinta años de antigüedad. Esta solución se enfocó en la administración integral del ciclo de vida del caso judicial, incluyendo el registro, seguimiento, intercambio de información y gestión documental asociada a cada proceso. La propuesta tecnológica buscó mejorar la interoperabilidad entre tribunales y otras entidades públicas, como la Policía de Victoria y organismos correccionales y de servicios sociales, permitiendo la captura e intercambio estructurado de información procesal. Asimismo, la implementación del portal CMS permitió a los usuarios externos realizar trámites en línea, como la presentación de documentos, autorización electrónica, seguimiento del estado del caso y gestión de pagos, contribuyendo a la reducción de tiempos administrativos y a una mayor trazabilidad de las actuaciones judiciales. Si bien estos resultados correspondieron a reportes institucionales del proyecto, su relevancia radica en demostrar cómo la automatización y la interoperabilidad pueden fortalecer la gestión de expedientes judiciales, constituyendo un referente aplicable a organismos como las procuradurías públicas en la defensa jurídica del Estado.

3.1.5. Sistemas de Gestión Procesal con capacidades de automatización avanzada en la Administración de Justicia. Yáñez (2023). España.

En el ámbito del Poder Judicial español, Yáñez (2023) describe el desarrollo de Sistemas de Gestión Procesal (SGP) implementados por Minsait, empresa del grupo Indra, en comunidades como la Junta de Andalucía y la Comunidad Valenciana. Estos sistemas actúan como núcleo operativo del ecosistema digital judicial, permitiendo la gestión integral del expediente electrónico, la clasificación automática de documentos, el control de plazos procesales y la trazabilidad de actuaciones. Asimismo, la plataforma Onesait Justice Platform incorpora inteligencia artificial y procesamiento de lenguaje

natural para analizar documentos jurídicos e identificar información relevante, reduciendo la carga administrativa en contextos de alto volumen documental. El autor concluye que la incorporación de automatización y análisis documental mejora la eficiencia y el seguimiento de los procesos judiciales, constituyendo un referente tecnológico aplicable a otras instituciones del sistema de justicia, como las procuradurías públicas.

3.1.6. Sistema de gestión procesal Avantius para la implementación del Expediente Judicial Electrónico en la Administración de Justicia. Bosco (2023). España.

En el proceso de modernización de la Administración de Justicia en España, el autor describe la implementación del sistema Avantius como una solución integral para la gestión del Expediente Judicial Electrónico (EJE) en comunidades autónomas como Navarra y el País Vasco (Euskadi). La plataforma fue validada por el Consejo General del Poder Judicial como sistema habilitado para la gestión procesal electrónica, consolidándose como herramienta central en la tramitación digital de expedientes. En Euskadi, el despliegue se realizó aproximadamente en un año, implicando la sustitución de sistemas previos y el diseño de una arquitectura web escalable orientada a mejorar la eficiencia operativa y reducir costos de mantenimiento. Asimismo, el proceso incluyó programas de capacitación dirigidos a miles de usuarios del sistema judicial, facilitando la adopción tecnológica y la estandarización de procesos. Finalmente, se señala que la implementación benefició directamente a unos 6.500 usuarios dentro de una comunidad profesional superior a 70.000 colaboradores, evidenciando el impacto positivo de la digitalización en la trazabilidad y gestión de expedientes judiciales, constituyendo un referente aplicable a otras instituciones públicas.

3.1.7. Sistema inteligente de asistencia jurídica basado en inteligencia artificial IURIX Mind. Unitech (2023).

La herramienta IURIX Mind fue desarrollada como una solución basada en inteligencia artificial generativa orientada a optimizar la gestión de expedientes judiciales electrónicos y el trabajo de profesionales del ámbito jurídico. Integrada con la plataforma IURIX Cloud Native, la herramienta incorpora capacidades avanzadas de procesamiento de lenguaje natural basadas en tecnología GPT de OpenAI, permitiendo a los usuarios interactuar mediante lenguaje natural para consultar información, analizar contenidos jurídicos y ejecutar acciones operativas dentro del sistema. Entre sus funcionalidades destacan la respuesta a preguntas sobre expedientes, la consulta de vencimientos, la presentación de escritos y la recepción de notificaciones, contribuyendo a agilizar tareas

administrativas y mejorar la toma de decisiones. El informe concluye que la integración de modelos de lenguaje con sistemas de gestión judicial permite aumentar la eficiencia operativa y constituye un avance relevante en la transformación digital de los servicios de justicia.

3.2. Herramientas utilizadas

En el desarrollo de las aplicaciones asociadas al sistema web SIE se empleó un conjunto de herramientas tecnológicas orientadas a la construcción de software, gestión de proyectos y despliegue en entornos productivos. En la capa backend se utilizó Python para la implementación de servicios y APIs, integrando librerías especializadas para el procesamiento de documentos y el uso de modelos de lenguaje (LLM) enfocados en la extracción y estructuración de información jurídica. Para la capa frontend se emplearon tecnologías web como HTML, CSS y JavaScript, junto con frameworks modernos que permitieron diseñar interfaces dinámicas, intuitivas y adaptadas a las necesidades de los usuarios.

Asimismo, se utilizaron herramientas de control de versiones como Git para la gestión del código fuente y la colaboración entre desarrolladores, además de entornos de desarrollo integrados (IDE) para la programación, pruebas y depuración del sistema. El despliegue y almacenamiento se realizaron mediante servicios de infraestructura en la nube, lo que permitió garantizar disponibilidad, escalabilidad y seguridad de la información procesada en el sistema.

Por otro lado, para la organización y gestión del proyecto se empleó la metodología ágil SCRUM, la cual facilitó la planificación iterativa del desarrollo mediante sprints, la definición de roles, la priorización de requerimientos y el seguimiento continuo del avance del proyecto. Este enfoque permitió una mejor adaptación a los cambios y una entrega progresiva de funcionalidades orientadas a optimizar la gestión de expedientes judiciales en la Procuraduría Pública del Ministerio del Interior del Perú

3.3. Casos de Éxito

3.3.1. Optimización del rendimiento en la extracción de información jurídica mediante arquitectura basada en microservicios y message brokers

El presente caso de éxito corresponde al estudio desarrollado por Ruiz et al. (2020), cuyo objetivo principal fue mejorar el rendimiento de los sistemas de extracción de información basados

en procesamiento de lenguaje natural (NLP) aplicados a documentos legales dentro de sistemas de gestión documental electrónica (EDMS). Este trabajo surge como respuesta a una problemática creciente en organizaciones públicas y privadas: el manejo masivo de documentos textuales, cuyo procesamiento manual resulta inviable debido al volumen, complejidad y especialización del lenguaje jurídico.

Desde el inicio, los autores identifican que la extracción de información en documentos legales es una tarea altamente demandante tanto en términos de tiempo como de recursos computacionales. Este tipo de documentos (contratos, escrituras, actos notariales, etc.) presentan características particulares como vocabulario técnico, gran extensión (hasta más de 100 páginas) y presencia de elementos complejos como firmas o sellos. Estas condiciones dificultan su análisis automático y hacen necesario el uso de sistemas avanzados de NLP. Sin embargo, estos sistemas suelen consumir grandes cantidades de CPU y memoria, afectando negativamente el rendimiento global de las plataformas donde se integran.

Frente a este contexto, los autores proponen una solución innovadora basada en la transformación de la arquitectura tradicional monolítica hacia una arquitectura distribuida basada en microservicios y el uso de un message broker. La idea central consiste en desacoplar el módulo de procesamiento NLP del sistema principal (EDMS) y convertirlo en un servicio independiente que pueda escalar horizontalmente.

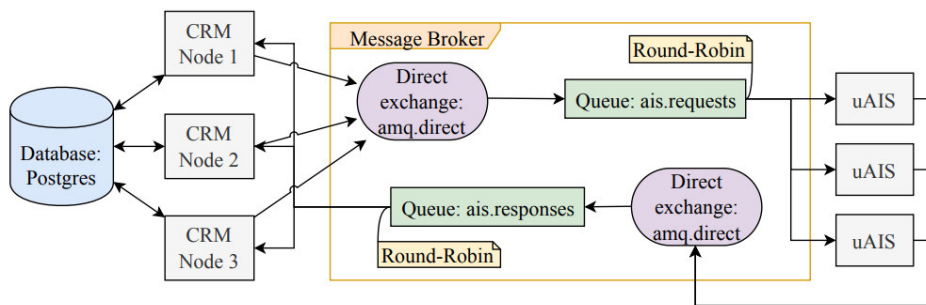
En la arquitectura tradicional (baseline), el sistema AIS (Analysis and Interpretation System) se encontraba integrado directamente dentro del sistema CRM llamado OnCustomer. Este sistema realizaba el procesamiento de documentos de forma secuencial, ejecutando tareas complejas como corrección ortográfica, lematización, parsing, reconocimiento de entidades (NER) y consultas ontológicas. Sin embargo, esta implementación presentaba múltiples limitaciones: procesamiento secuencial, escalabilidad únicamente vertical, alto consumo de recursos en cada nodo del sistema, ejecución periódica (no en tiempo real) y falta de modularidad en las etapas del procesamiento.

Para superar estas limitaciones, los autores diseñan una arquitectura basada en Message Oriented Middleware (MOM). En este nuevo enfoque, el sistema se divide en tres componentes principales:

- **EDMS (CRM OnCustomer):** encargado de recibir solicitudes de procesamiento de documentos.
- **Microservicios NLP (uAIS):** encargados de ejecutar las tareas de extracción de información.
- **Message Broker (RabbitMQ):** encargado de gestionar la comunicación entre los componentes mediante colas de mensajes.

Figura 14

Arquitectura del sistema de comunicación propuesto



Nota. A la izquierda se encuentran las instancias de CRM que realizan solicitudes a los servicios uAIS ubicados a la derecha. La gestión de estas solicitudes se realiza a través del Agente de Mensajes. Basado en: Ruiz et al., 2020

El flujo de trabajo funciona de la siguiente manera: cuando un usuario solicita el procesamiento de un documento, el sistema EDMS envía una solicitud al message broker. Este la coloca en una cola y la distribuye entre los microservicios disponibles utilizando un esquema de balanceo de carga (Round-Robin). Cada microservicio procesa el documento de forma independiente y devuelve el resultado al broker, el cual lo reenvía al sistema original. Este enfoque permite que múltiples documentos sean procesados en paralelo, mejorando significativamente la eficiencia del sistema. Uno de los aspectos clave de la implementación fue el uso de RabbitMQ como broker de mensajería, debido a sus ventajas como persistencia de mensajes, alta disponibilidad, soporte de múltiples protocolos, bajo consumo de memoria y facilidad de integración. Además, el uso de microservicios permite replicar instancias del módulo NLP según la demanda, logrando una escalabilidad horizontal eficiente.

Para validar la propuesta, los autores realizaron experimentos utilizando documentos reales del sistema OnCustomer, específicamente contratos de compraventa de entre 75 y 100 páginas. El

sistema debía extraer aproximadamente 55 tipos de datos por documento, incluyendo información de personas, propiedades y transacciones. Se implementaron ocho microservicios uAIS en paralelo para procesar las solicitudes. Los resultados obtenidos evidencian una mejora significativa en el rendimiento del sistema. En términos de tiempo de procesamiento, se logró una reducción del 77%, pasando de aproximadamente 450 segundos a menos de 100 segundos para procesar 200 documentos. Asimismo, se observó un mejor aprovechamiento del CPU, ya que el sistema dejó de operar de forma pasiva (esperando resultados) y pasó a utilizar plenamente los recursos disponibles.

3.3.2. Clasificación de textos legales mediante el modelo BERT-MS

El estudio de Zhong et al. (2020) aborda el problema de la clasificación automática de textos en el ámbito jurídico, el cual se ha vuelto cada vez más relevante debido al crecimiento exponencial de documentos legales digitales. En contextos como tribunales, estudios de abogados y entidades financieras, la correcta categorización de documentos resulta fundamental para facilitar la búsqueda de información, mejorar la eficiencia operativa y apoyar la toma de decisiones. Sin embargo, esta tarea presenta una alta complejidad debido a las características propias del lenguaje jurídico, el cual se distingue por su formalidad, terminología especializada y estructuras sintácticas complejas. En este sentido, los autores identifican que los enfoques tradicionales de procesamiento de lenguaje natural no logran capturar adecuadamente las relaciones semánticas profundas presentes en este tipo de textos, lo que limita su desempeño en tareas de clasificación.

A partir de esta problemática, la investigación propone un enfoque basado en modelos avanzados de aprendizaje profundo, específicamente en arquitecturas tipo Transformer. En particular, se parte del uso del modelo BERT, el cual ha demostrado una gran capacidad para comprender el contexto del lenguaje gracias a su mecanismo de atención bidireccional. No obstante, los autores señalan que, aunque BERT representa un avance significativo frente a modelos anteriores, aún presenta limitaciones cuando se aplica a documentos legales extensos, ya que no siempre logra capturar de manera eficiente la información más relevante a nivel global. Por ello, el estudio introduce una variante mejorada denominada BERT-MS, cuyo objetivo es optimizar la representación semántica de los textos y mejorar el rendimiento en tareas de clasificación dentro del dominio legal.

Para validar la propuesta, los autores desarrollan una metodología experimental basada en la comparación de múltiples modelos bajo condiciones controladas. En primer lugar, se realiza un

proceso de preprocesamiento de los datos, el cual incluye la limpieza de los textos, la eliminación de elementos irrelevantes y la normalización del contenido. Posteriormente, los documentos son transformados en representaciones que pueden ser procesadas por los modelos mediante técnicas de tokenización. Una vez preparados los datos, se procede al entrenamiento de distintos modelos de clasificación, abarcando tanto enfoques tradicionales como modelos más avanzados basados en aprendizaje profundo. Finalmente, se evalúa el desempeño de cada modelo utilizando métricas estándar que permiten medir tanto el rendimiento global como la capacidad de generalización entre diferentes clases.

El conjunto de datos utilizado en el estudio está compuesto por documentos legales reales, organizados en tres categorías principales: casos de divorcio, casos laborales y casos relacionados con préstamos. Esta diversidad de categorías permite evaluar la capacidad de los modelos para distinguir entre diferentes tipos de contenido jurídico, cada uno con características particulares. Los autores aseguran que el dataset es representativo del dominio, lo que resulta clave para garantizar la validez de los resultados. Asimismo, los datos son divididos en conjuntos de entrenamiento y prueba, lo que permite evaluar el desempeño de los modelos en escenarios controlados y evitar problemas de sobreajuste.

En cuanto a los modelos evaluados, el estudio incluye una amplia variedad de enfoques. Por un lado, se consideran modelos tradicionales como TextCNN, DPCNN y LSTM, los cuales han sido ampliamente utilizados en tareas de clasificación de texto. Estos modelos se basan en la extracción de características locales o en el modelado de secuencias, pero presentan limitaciones para capturar dependencias de largo alcance. Por otro lado, se incorpora el modelo BiDAF, el cual introduce mecanismos de atención que permiten una mejor representación del contexto. Finalmente, se evalúan modelos basados en Transformers, incluyendo BERT y la propuesta BERT-MS. Este último constituye la principal contribución del estudio, ya que incorpora mejoras orientadas a capturar de manera más efectiva la información relevante del documento completo.

Tabla 1*Resultados experimentales sobre la extracción de elementos.*

	Divorce		Labor		Loan	
Model	MiF	MaF	MiF	MaF	MiF	MaF
TextCNN	78.7	65.9	76.4	54.4	80.3	60.6
DPCNN	81.3	64.0	79.8	47.4	81.4	42.5
LSTM	80.6	67.3	81.0	52.9	80.4	53.1
BiDAF	83.1	68.7	81.5	59.4	80.5	63.1
BERT	83.3	69.6	76.8	43.7	78.6	39.5
BERT-MS	84.9	72.7	79.7	54.5	81.9	64.1

Nota. Aquí, MiF y MaF denotan micro-F1 y macro-F1, respectivamente. Basado en: Zhong et al., 2020

Los resultados obtenidos evidencian diferencias significativas en el desempeño de los modelos. Para la evaluación se utilizan las métricas Micro-F1 y Macro-F1, las cuales permiten analizar tanto el rendimiento global como el equilibrio entre clases. En la categoría de divorcio, el modelo BERT-MS alcanza los mejores resultados, con un valor de 84.9 en Micro-F1 y 72.7 en Macro-F1, superando a todos los modelos comparados. En la categoría laboral, el modelo BiDAF obtiene el mayor desempeño en ciertas métricas, lo que sugiere que los mecanismos de atención siguen siendo altamente competitivos en determinados contextos. Sin embargo, BERT-MS mantiene un rendimiento sólido, demostrando su capacidad de generalización. En la categoría de préstamos, BERT-MS vuelve a posicionarse como el mejor modelo, alcanzando valores de 81.9 en Micro-F1 y 64.1 en Macro-F1, lo que confirma su superioridad en múltiples escenarios.

El análisis de estos resultados permite concluir que los modelos basados en Transformers, y en particular la propuesta BERT-MS, logran superar las limitaciones de los enfoques tradicionales. Esto se debe a su capacidad para capturar relaciones semánticas complejas y representar de manera más completa el contexto del texto. Además, la mejora introducida en BERT-MS demuestra que es posible adaptar modelos generales a dominios específicos, obteniendo un desempeño superior en tareas especializadas. Este aspecto resulta especialmente relevante en el ámbito jurídico, donde la precisión y la comprensión del contexto son fundamentales.

3.3.3. Implementación de redes neuronales convolucionales para la clasificación automática multietiqueta de documentos legales

El estudio realizado por Ming et al. (2020) presenta una investigación centrada en la aplicación de técnicas avanzadas de procesamiento de lenguaje natural para mejorar la comprensión y el tratamiento automático de textos dentro de un dominio específico. Los autores parten de la necesidad creciente de gestionar grandes volúmenes de información textual, lo cual representa un desafío tanto para organizaciones como para sistemas automatizados. En este contexto, se destaca que los métodos tradicionales de análisis de texto resultan insuficientes para capturar la complejidad semántica y contextual presente en documentos especializados, lo que motiva la exploración de modelos más sofisticados basados en aprendizaje profundo. De esta manera, el trabajo se enfoca en proponer y evaluar un enfoque que permita mejorar significativamente el rendimiento en tareas de procesamiento de texto, contribuyendo al avance del estado del arte en este campo.

En la introducción del estudio, los autores realizan una revisión del problema y contextualizan la importancia de desarrollar modelos capaces de comprender el lenguaje de manera más precisa. Se menciona que, en aplicaciones reales, como sistemas de recomendación, análisis de documentos o clasificación de información, la calidad de la representación del texto es un factor determinante. En este sentido, se reconoce el impacto que han tenido los modelos basados en arquitecturas neuronales profundas, especialmente aquellos que incorporan mecanismos de atención, los cuales permiten identificar relaciones relevantes dentro del texto sin depender estrictamente de la secuencia lineal. Esta capacidad representa un cambio significativo respecto a modelos previos, como redes recurrentes o convolucionales, que presentan limitaciones al manejar dependencias de largo alcance.

A nivel metodológico, el trabajo propone un enfoque estructurado que combina distintas etapas para el procesamiento y análisis de los datos. En primer lugar, se realiza la preparación del conjunto de datos, lo cual implica la recolección, limpieza y organización de los textos. Este proceso es fundamental para garantizar la calidad de la información utilizada en el entrenamiento de los modelos. Posteriormente, se lleva a cabo la transformación de los datos en representaciones numéricas, utilizando técnicas de tokenización y codificación que permiten a los modelos interpretar el lenguaje natural. Una vez completada esta etapa, se procede al diseño y entrenamiento de los

modelos propuestos, los cuales se basan en arquitecturas modernas que aprovechan mecanismos de atención para mejorar la comprensión del contexto.

El conjunto de datos utilizado en la investigación está cuidadosamente seleccionado para reflejar las características del problema abordado. Los autores aseguran que los datos son representativos del dominio de estudio, lo que permite evaluar de manera adecuada el desempeño de los modelos en escenarios reales. Además, el dataset es dividido en diferentes subconjuntos para entrenamiento, validación y prueba, lo cual permite medir la capacidad de generalización de los modelos y evitar problemas como el sobreajuste. Este enfoque experimental garantiza que los resultados obtenidos sean confiables y puedan ser comparados de manera objetiva con otros métodos existentes.

En cuanto al modelo propuesto, los autores destacan el uso de una arquitectura basada en mecanismos de atención, la cual permite capturar relaciones complejas dentro del texto. A diferencia de enfoques tradicionales, este tipo de modelo no procesa la información de manera estrictamente secuencial, sino que analiza el contexto completo de cada palabra en relación con el resto del texto. Esta característica resulta especialmente útil en documentos largos o complejos, donde la información relevante puede estar distribuida en distintas partes del contenido. Asimismo, el modelo incorpora mejoras que optimizan su capacidad de representación, permitiendo obtener resultados más precisos en las tareas evaluadas.

Para validar la efectividad del enfoque propuesto, los autores realizan una serie de experimentos comparativos con distintos modelos de referencia. Estos incluyen tanto métodos tradicionales como modelos más avanzados, lo que permite establecer una comparación amplia del desempeño. La evaluación se lleva a cabo utilizando métricas estándar que permiten medir la precisión y la calidad de las predicciones. A través de estos experimentos, se busca demostrar que el modelo propuesto no solo es competitivo, sino que ofrece mejoras significativas respecto a los enfoques existentes.

Tabla 2

Tabla comparativa de rendimiento de diferentes modelos.

	Model	Precision	Recall	F1
Binary relevance	LR	0.6132	0.6035	0.6083
	KNN	0.5821	0.5704	0.5762
	GaussianNB	0.4914	0.4954	0.4934
Classifier chains	LR	0.6573	0.6641	0.6607
	KNN	0.6132	0.6039	0.6085
	GaussianNB	0.6398	0.6232	0.6314
Label powerset	LR	0.6291	0.6001	0.6143
	KNN	0.6131	0.5917	0.6022
	GaussianNB	0.4254	0.4396	0.4324
fastText		0.843	0.8434	0.8432
MLKNN		0.6087	0.6025	0.6056
MLTCNN		0.9807	0.9961	0.9884

Nota. Se presentan diferentes modelos de clasificación de texto, evaluados bajo las métricas de Precisión (Precisión), Exhaustividad (Recall) y Valor F1. Basado en: Ming et al., 2020

Los resultados obtenidos evidencian que el modelo propuesto logra un desempeño superior en comparación con otros métodos evaluados. En particular, se observa una mejora consistente en las métricas utilizadas, lo que indica una mayor capacidad para comprender el contexto y realizar predicciones precisas. Este resultado confirma la hipótesis planteada por los autores, quienes sostienen que la incorporación de mecanismos de atención y arquitecturas avanzadas permite superar las limitaciones de los enfoques tradicionales. Además, se destaca que el modelo mantiene un buen nivel de generalización, lo cual es fundamental para su aplicación en escenarios reales.

El análisis de los resultados también permite identificar algunos aspectos relevantes sobre el comportamiento de los modelos. Por ejemplo, se observa que los modelos tradicionales tienden a presentar un rendimiento inferior cuando se enfrentan a textos complejos o con estructuras poco convencionales. En contraste, los modelos basados en atención logran adaptarse mejor a estas condiciones, lo que les permite mantener un desempeño más estable. Asimismo, los autores señalan que la calidad del dataset y el proceso de preprocesamiento tienen un impacto significativo en los resultados, lo que resalta la importancia de estas etapas dentro del flujo de trabajo.

3.3.4. Extracción inteligente de información legal con BERT-BiLSTM-CRF

El estudio desarrollado por Gu et al. (2020) aborda una problemática crítica en la modernización de los sistemas judiciales: la necesidad de extraer información estructurada de documentos legales de manera automática. El reconocimiento de entidades nombradas (NER) en este campo es un desafío mayor debido a que la industria judicial es intensiva en conocimiento y genera volúmenes masivos de datos, como documentos de sentencias y casos guía, que contienen terminología técnica extremadamente compleja. Los métodos tradicionales de aprendizaje automático y las reglas manuales suelen fallar en este entorno porque dependen de la selección de características por parte de expertos humanos, lo que limita su capacidad de generalización y robustez ante la ambigüedad del lenguaje legal. Para resolver esto, los autores proponen una arquitectura de vanguardia denominada BERT-BiLSTM-CRF, la cual elimina la necesidad de intervención humana en el diseño de características y permite que el sistema aprenda directamente de la semántica profunda de los textos.

La base de este éxito radica en la integración del modelo de lenguaje pre-entrenado BERT (Bidirectional Encoder Representations from Transformers), que revoluciona la forma en que se representan las palabras. A diferencia de los vectores de palabras estáticos como Word2Vec, que asignan un único valor a una palabra sin importar su contexto, BERT utiliza un codificador de transformador bidireccional que analiza la oración completa simultáneamente. Esto es fundamental en el ámbito jurídico, donde una palabra puede tener significados radicalmente distintos según el tipo de delito o el procedimiento civil en cuestión; BERT genera representaciones vectoriales dinámicas que capturan estas sutilezas semánticas de forma global.

Una vez que BERT ha transformado el texto en vectores enriquecidos, esta información se procesa a través de una capa de memoria a corto y largo plazo bidireccional (BiLSTM). El modelo LSTM es particularmente eficaz para el campo judicial porque utiliza un sistema de "puertas" — entrada, olvido y salida— que le permite identificar qué fragmentos de una oración larga son relevantes y cuáles deben descartarse, resolviendo el problema de las dependencias a largo plazo en textos legales extensos. Al procesar la información en ambas direcciones (de principio a fin y viceversa), la BiLSTM logra una codificación semántica superior que prepara los datos para la etapa final de etiquetado.

El último componente de esta arquitectura es la capa de Campo Aleatorio Condicional (CRF), que actúa como un filtro de coherencia para las etiquetas predichas. En el reconocimiento de entidades, existen dependencias lógicas estrictas; por ejemplo, una etiqueta que indica el "interior" de una entidad no puede seguir lógicamente a una que indica el "exterior" de cualquier entidad. El CRF analiza las relaciones entre etiquetas adyacentes y calcula la secuencia de etiquetas globalmente óptima, asegurando que la estructura final de la entidad reconocida (como un "sujeto legal" o un "hecho legal") sea gramatical y lógicamente válida.

Para validar este enfoque, los investigadores compilaron un corpus judicial robusto utilizando dos fuentes principales: por un lado, 294 artículos de sitios web oficiales de la Corte Suprema y procuradurías de China y, por otro, el diccionario "Compilación de la Ley Vigente de China", sumando millones de palabras en términos legales y disposiciones constitucionales. Utilizando una estrategia de etiquetado manual supervisada por cinco expertos para resolver controversias, el equipo clasificó las entidades en cuatro categorías clave: documentos legales (DOC), sujetos legales (SUB), objetos legales (OBJ) y hechos legales (FAC), siendo esta última la más predominante con un 58.6% del corpus.

Tabla 3

Resultados de Diferentes Modelos al Reconocimiento de Entidades

Models	P	R	F1
BiLSTM	79.28	74.20	78.48
BiLSTM-CRF	82.85	81.12	81.59
Radical-BiLSTM-CRF	85.16	84.62	84.95
BERT-BiLSTM-CRF	88.86	87.49	87.97

Nota. Se muestran los resultados en temas de rendimiento de cada uno de los modelos. Basado en: Gu et al., 2020

Los resultados experimentales demostraron de manera contundente la superioridad de esta propuesta técnica. El modelo BERT-BiLSTM-CRF alcanzó una precisión del 88.86% y un valor F1 del 87.97%, superando significativamente a modelos estándar como BiLSTM-CRF (81.59%) y Radical-BiLSTM-CRF (84.95%). Este incremento del 6.42% en el valor F1 respecto a los modelos que no usan BERT confirma que la capacidad de extracción de características de este modelo

preentrenado es excepcionalmente fuerte para el lenguaje técnico. El estudio concluye que este sistema no solo resuelve el problema de la ambigüedad de las palabras, sino que también ofrece una herramienta altamente robusta y generalizable para asistir a los profesionales del derecho en la gestión inteligente de casos y la reducción de la carga de trabajo manual.

3.3.5. Sistema Juco-IS para gestión judicial digital

El proyecto, denominado JuCo-IS por Acoba et al. (2020), representa un avance significativo en la modernización del sistema judicial de Filipinas, específicamente diseñado para implementarse en los Tribunales Regionales de Primera Instancia (RTC). Esta iniciativa surge como respuesta a la necesidad de cumplir con los objetivos nacionales de e-gobierno, los cuales buscan crear una nación digitalmente empoderada que ofrezca servicios transparentes y centrados en el ciudadano. Históricamente, la gestión de documentos en los tribunales ha presentado dificultades críticas para archiveros y administradores, lo que motivó el desarrollo de una estructura basada en la web que permitiera digitalizar los procesos legales desde el registro de un caso hasta las audiencias y la actualización de expedientes. El éxito de este sistema se fundamenta en la integración de dos conceptos pilares: el Sistema de Gestión de Información de Casos (CIMS) y el Registro Electrónico de Registros Judiciales (ECRK), los cuales sirven de base para optimizar el rol de las tecnologías de la información en la gobernanza.

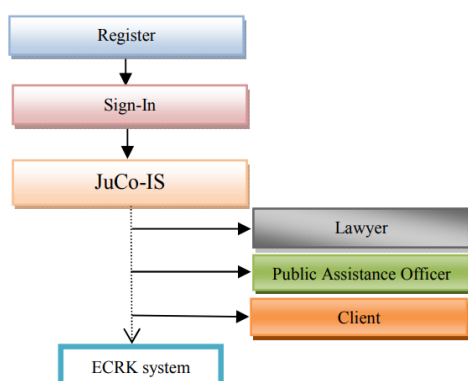
Desde una perspectiva técnica, el desarrollo del JuCo-IS siguió un Ciclo de Vida de Desarrollo de Sistemas (SDLC) que comprendió las fases de planificación, análisis, diseño y evaluación. Para garantizar la robustez de la plataforma, los investigadores utilizaron métodos experimentales y descriptivos con el fin de determinar los requisitos funcionales necesarios para un sistema de registro electrónico eficiente sobre una conexión a internet. La arquitectura del software se diseñó mediante diversos escenarios de casos de uso y diagramas de clases que definen con precisión los atributos y las relaciones entre las entidades del tribunal y los expedientes. Este enfoque técnico permitió que el sistema no solo procese información, sino que también facilite el intercambio de documentos y el almacenamiento seguro de datos, eliminando progresivamente el uso de papel en las oficinas judiciales.

El funcionamiento operativo del JuCo-IS se divide estratégicamente para atender a diferentes tipos de actores, clasificados como personal de la corte y usuarios externos. El personal judicial, que

incluye a los oficiales de asistencia pública, utiliza el sistema para gestionar el flujo de los procedimientos, organizar cronogramas de audiencias y redactar documentos finales. Por otro lado, los usuarios externos, principalmente abogados y ciudadanos (clientes), acceden a un portal web que les permite registrar casos, presentar peticiones de audiencia y solicitar documentos judiciales de manera remota. Esta interacción digital se facilita mediante una interfaz gráfica de usuario (GUI) diseñada para ser amigable, lo que reduce la curva de aprendizaje para los principiantes y mejora la experiencia general del usuario al navegar por el sistema.

Figura 15

Diagrama del Sistema Juco-IS



Nota. Se muestra el modelado del Sistema Juco-IS a nivel de bloques. Basado en: Acoba et al., 2020

Un componente crítico para el éxito de esta implementación es la infraestructura de red y la seguridad de la información. El JuCo-IS está diseñado para operar en entornos de Red de Área Local (LAN) y Red de Área Amplia (WAN), permitiendo una expansión directa según las necesidades de crecimiento del tribunal. Para garantizar la integridad de la justicia, el sistema cuenta con disposiciones estrictas para realizar copias de seguridad de los archivos digitales y de la base de datos en sitios de respaldo especificados, protegiendo los registros contra posibles pérdidas. Además, el proceso de evaluación del sistema incluyó pruebas de usabilidad exhaustivas para validar que la interacción entre el ingreso de datos y el sistema de registro cumpliera con los estándares de exactitud requeridos en el ámbito legal.

En conclusión, la implementación del JuCo-IS ha demostrado ser una herramienta de gran ayuda para la sociedad al digitalizar integralmente el ciclo de vida de un caso judicial. Los resultados

del proyecto permiten a los administradores autorizados supervisar una amplia gama de reportes, incluyendo estados de los casos, actualizaciones recientes y listas de clientes, variando según la duración de los datos presentados. Al optimizar el flujo de trabajo de los oficiales judiciales y abogados, el sistema no solo incrementa la eficiencia operativa y la transparencia, sino que también proporciona un modelo escalable que los investigadores planean integrar en tribunales de jerarquías superiores en el futuro. Este caso de éxito reafirma que el uso estratégico de sistemas de información basados en la web es el camino para descongestionar el poder judicial y ofrecer una justicia más ágil y accesible para todos los ciudadanos

3.3.6. Extracción de Hechos Judiciales mediante Aprendizaje de Patrones semi supervisados e Integración de Ontologías

Para este caso de éxito, desarrollado por Anu y Sangeetha (2021) de la National Institute of Technology en India, presenta un enfoque innovador para la extracción de hechos en documentos judiciales mediante el uso de aprendizaje de patrones semi supervisados integrado con ontologías. La introducción de este estudio destaca que, ante el crecimiento masivo de sentencias electrónicas, es imperativo contar con sistemas de extracción de información (IE) que identifiquen no solo entidades, sino los roles específicos que estas cumplen en un proceso legal. A diferencia de los métodos tradicionales que dependen de reglas manuales rígidas, esta propuesta utiliza una Ontología de Casos Judiciales (JCO) para guiar el proceso de aprendizaje (bootstrapping), permitiendo que el sistema aprenda patrones de extracción complejos a partir de un conjunto muy pequeño de hechos semilla.

La problemática que motiva esta investigación reside en la naturaleza intrínseca del dominio judicial, el cual es considerado un "dominio cerrado" donde no abundan los conjuntos de datos etiquetados para entrenar modelos de aprendizaje profundo o supervisado. Los autores señalan que generar estos datos manualmente requiere una participación humana exhaustiva y costosa, además de que los sistemas basados únicamente en reglas manuales son difíciles de mantener y no son fácilmente portables a otros idiomas o jurisdicciones. Asimismo, los métodos convencionales de vectores de palabras suelen fallar al capturar la polisemia o ambigüedad de los términos legales, lo que genera la necesidad de un modelo que integre el conocimiento del dominio de forma explícita para distinguir, por ejemplo, entre un peticionario, un juez o un testigo dentro de la misma narrativa procesal.

La metodología propuesta se estructura en tres fases secuenciales: preprocesamiento, aprendizaje y extracción. En la fase inicial, se realiza una normalización del texto para resolver abreviaturas legales (como transformar "PW" en "primary witness") y se aplican etiquetas de entidades nombradas y conceptos ontológicos basados en la JCO. Durante la fase de aprendizaje, el sistema utiliza los hechos semilla para identificar contextos en el corpus de entrenamiento y generar plantillas de patrones de forma automática. Estos patrones son evaluados mediante una métrica de "fuerza" para seleccionar solo aquellos que son predictores confiables de cada categoría de hecho. Finalmente, en la fase de extracción, estos patrones validados se aplican sobre el conjunto de datos de prueba para identificar automáticamente 12 clases de hechos judiciales, incluyendo números de caso, nombres de abogados, investigadores y la decisión final del juez.

Tabla 4

Rendimiento del enfoque propuesto en el conjunto de datos judiciales

Facts	Precision	Recall	F-score
Case No.	100	100	100
Petitioner	100	100	100
Respondent	100	100	100
Judge	100	100	100
Attorney	100	100	100
Investigator	100	100	100
Witness	100	96	97.95
FIR	96.11	97.05	96.6
Cases_referred	96.2	98.6	97.4
Rules	100	88	93.6
Judgement_date	100	100	100
Final_decision	100	98	98.9

Nota. Se muestran las métricas de rendimiento del enfoque propuesto para la extracción de hechos en el dominio judicial, evaluando 12 categorías diferentes mediante precisión, exhaustividad (recall) y valor F1.

Basado en: Anu y Sangeetha, 2021

La implementación y los resultados de este caso de éxito demuestran una eficacia superior a los métodos existentes, logrando una precisión promedio del 99.35% y una exhaustividad (recall) del 98.13%. El sistema fue probado en un corpus de 500 documentos judiciales relacionados con casos criminales, logrando extraer con una precisión del 100% categorías críticas como el peticionario, el demandado, el juez y la fecha de la sentencia. Aunque se detectaron ligeros retos en la identificación de testigos debido a problemas de detección de límites de las palabras, el sistema mostró una robustez

notable al funcionar sin necesidad de expresiones regulares generadas manualmente o conjuntos de datos anotados masivos. En última instancia, la gran ventaja de este modelo es su portabilidad, ya que puede adaptarse a otros dominios como el médico o financiero simplemente sustituyendo la ontología de base, lo que lo convierte en un pilar para el desarrollo de sistemas expertos modernos.

3.3.7. Extracción de Entidades en Documentos de Sentencias Judiciales Mediante el Método de Reglas de Expresiones Regulares

La investigación desarrollada por Kainan y Xin (2021) de la Universidad de Seguridad Pública de China, presenta una investigación enfocada en la extracción de entidades de lenguaje judicial utilizando un método basado en reglas de expresiones regulares. La introducción del estudio destaca que, en la actual era del Estado de Derecho e inteligencia, la tecnología de Procesamiento de Lenguaje Natural (PLN) juega un papel fundamental para transformar la vasta cantidad de textos judiciales no estructurados en información profesional aprovechable. El reconocimiento de entidades nombradas (NER) se define aquí como el proceso automático de identificar nombres de entidades conocidos o definidos por el ser humano a partir de reglas de práctica histórica. Los autores eligen como objeto de estudio los documentos de sentencias judiciales debido a que poseen un formato estándar, están finamente redactados y son de fácil procesamiento en comparación con otros textos no estructurados.

La problemática identificada por los investigadores radica en que la mayoría de los estudios de NER se especializan en campos generales, dejando una brecha de exploración en el ámbito judicial. Actualmente, el procesamiento y archivo de una gran cantidad de textos judiciales todavía depende del trabajo manual, lo cual es ineficiente ante el crecimiento explosivo del número de sentencias. Aunque existen métodos de aprendizaje profundo como BERT o CNN, estos requieren conjuntos de datos anotados a gran escala y un soporte computacional de altos recursos, lo que plantea desafíos significativos para el campo judicial donde la disponibilidad de datos anotados es limitada. Además, el lenguaje judicial difiere del lenguaje cotidiano, lo que genera problemas de comprensión pública si no se estandariza y extrae correctamente la información de los órganos judiciales.

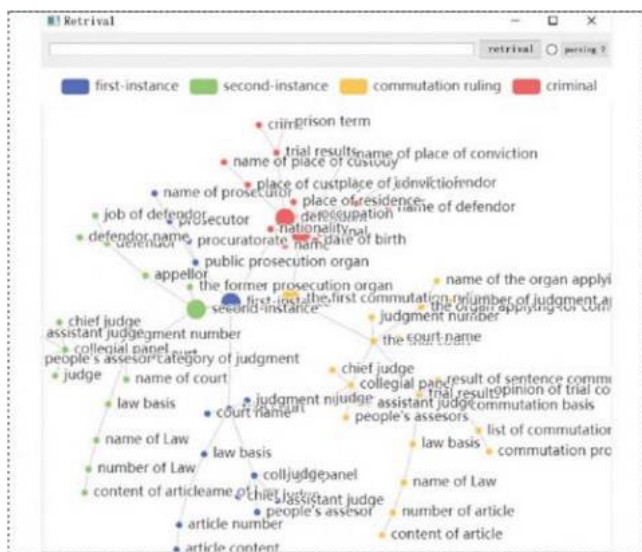
La metodología propuesta se centra en un método basado en reglas de expresiones regulares, seleccionado tras contrastar su efectividad y costo frente a herramientas como GATE y algoritmos de aprendizaje automático. El flujo de investigación comienza con la obtención de sentencias mediante el marco de rastreo "scrapy" desde el sitio web China Judgments Online. Las entidades se dividen en

términos regulares (cortos y con estructura distintiva, como nombres de personas o instituciones) y términos irregulares (narrativas largas sobre el proceso del juicio). Los autores analizan la estructura de las sentencias penales de primera instancia, que se componen principalmente de siete partes fijas, lo que facilita el diseño de reglas específicas para extraer desde el nombre del tribunal hasta la información de los oficiales judiciales

En cuanto a la implementación, el sistema extrae los juicios línea por línea y utiliza el formato JSON para encapsular los resultados del reconocimiento en el back-end, facilitando la interacción con el front-end. Al parsear una sola sentencia, se inicializa un diccionario de marco que se va completando o reemplazando con los valores extraídos durante el proceso. Las primeras líneas del documento se utilizan para extraer el nombre de la corte y el número de sentencia, seguidas por la información personal del demandado, el lugar de custodia y la información del defensor. Un aspecto clave de la implementación es la extracción de la base legal de la decisión, donde el sistema identifica secuencialmente el nombre de la ley y los artículos o sub-artículos correspondientes citados en el caso.

Figura 16

Mapa de Circulación de Entidades



Nota. Se representa un Mapa de Circulación de Entidades generado automáticamente a partir de sentencias judiciales. Basado en: Kainan y Xin, 2021

Los resultados de este caso de éxito se visualizan a través de un "mapa de circulación de entidades". Este mapa muestra de forma detallada la transferencia de información y las relaciones entre personas (acusado, apelante, criminal) y organizaciones (tribunales, fiscalías, prisiones) en las diferentes etapas del proceso judicial. Además, las entidades extraídas se integran en un grafo de conocimiento, permitiendo excavar relaciones implícitas, como la conexión entre delincuentes que comparten el mismo lugar de residencia para calcular la relación entre castigo y geografía. El estudio concluye que el uso de expresiones regulares es el método más adecuado para este escenario debido a la alta regularidad de los documentos y la falta de un corpus anotado suficiente para métodos de aprendizaje automático.

3.3.8. "Desarrollo del Software 'Radicapp' para la Gestión de Información y Mitigación de la Congestión en los Despachos Judiciales de Colombia"

Este caso de éxito, desarrollado por Rojas et al. (2022) de la Universidad Cooperativa de Colombia, se centra en el diseño de una aplicación informática como estrategia para mitigar la congestión judicial en el país. La investigación propone el uso de herramientas tecnológicas de gestión de información, específicamente software web, para optimizar las labores administrativas de los despachos judiciales colombianos. El estudio toma como referencia particular al Juzgado 01 Civil Circuito de Fusagasugá, Cundinamarca, buscando demostrar cómo la automatización del flujo documental puede reducir significativamente los tiempos de atención al usuario y mejorar la calidad de los reportes estadísticos exigidos por la Rama Judicial.

La problemática central identificada es el retraso sistémico y la congestión en los despachos judiciales, causada en gran parte por la persistencia de métodos manuales para el registro y seguimiento de procesos. Actualmente, los empleados deben anotar avances en libros físicos, lo que genera un alto margen de error y hace que identificar el estado de un caso sea una labor dispendiosa y, a veces, casi imposible. Además, Colombia presenta un déficit significativo de jueces, con solo 12.2 por cada 100,000 habitantes frente al estándar internacional de 65, lo que agrava la carga operativa y dificulta el cumplimiento de la Ley de Transparencia sobre el acceso a la información pública nacional.

Tabla 5

Etapas en el desarrollo del software de gestión de información

Etapas	Descripción
Primera	Definir las necesidades de investigación, se realiza un análisis y conocimiento a fondo del objeto de estudio, con el propósito de realizar un diagnóstico del comportamiento y actividades que se llevan a cabo dentro del juzgado, referente a la gestión de información.
Segunda	Elaborar los requisitos funcionales y no funcionales necesarios para el diseño del prototipo del software de gestión; además, se clasifica la información de acuerdo con el diagnóstico de las necesidades de la investigación.
Tercera	Se realizan los modelos de la arquitectura del software, tomando como insumo los requisitos funcionales y no funcionales recopilados.
Cuarta	Diseño y desarrollo del software de prueba, que se considera el prototipo, y evaluación del cumplimiento de los requerimientos por parte del personal del juzgado.
Quinta	Instalación del software en los equipos de cómputo del juzgado y ejecución del plan de pruebas de funcionamiento.
Sexta	Entrega final del software al juzgado.

Nota. Se detallan las seis etapas fundamentales seguidas para el diseño, desarrollo e implementación del software de gestión de información judicial en el despacho objeto de estudio. Basado en: Rojas et al., 2022

La metodología empleada para abordar este desafío es de tipo evaluativa con un método deductivo preexperimental de caso único. El proceso investigativo se estructuró en seis etapas fundamentales que iniciaron con un diagnóstico profundo del comportamiento del juzgado y la definición de requisitos funcionales y no funcionales para el software. Los investigadores analizaron diversas herramientas de código abierto existentes, como Alfresco y Nuxeo, pero concluyeron que se requería un aplicativo diseñado específicamente para las fases del expediente judicial colombiano y la interacción entre múltiples roles (jueces, abogados y demandantes).

La implementación se materializó a través del desarrollo de la aplicación web denominada www.radicapp.com, construida con lenguajes y entornos de trabajo de código abierto como PHP y el framework Laravel. El sistema utiliza una arquitectura de cloud hosting para garantizar la disponibilidad del servicio sin interrupciones y una base de datos relacional MySQL para el almacenamiento flexible de alto rendimiento. El software cuenta con módulos específicos para radicación, actuaciones, memoriales y un componente estadístico que automatiza la generación de informes requeridos por el Sistema Estadístico de la Rama Judicial (Sierju).

Los resultados obtenidos validan la eficacia de la herramienta al transformar radicalmente la gestión del conocimiento dentro del despacho judicial. La implementación permitió una agilización drástica en la radicación y seguimiento de procesos, eliminando los cuellos de botella del registro

manual y garantizando la veracidad de los datos estadísticos reportados. El sistema recibió una alta receptividad por parte de los usuarios externos, como abogados, debido a su diseño intuitivo y a la posibilidad de consultar el historial de actuaciones en tiempo real. Finalmente, el estudio destaca que radicapp.com tiene el potencial de ser escalado masivamente a otros despachos del país, contribuyendo a la construcción de un sistema judicial más inclusivo y eficiente.

3.3.9. Sistema web basado en metodología XP para mejorar la gestión de expedientes en el estudio jurídico Gamión Asociados

El proyecto desarrollado por Matos (2022), tuvo como objetivo principal mejorar la gestión de expedientes mediante la implementación de una solución tecnológica. La investigación se originó a partir de la problemática identificada en el estudio jurídico Gamión Asociados, donde los procesos de registro, seguimiento y control de expedientes se realizaban de manera manual, generando demoras, pérdida de información, dificultades en la búsqueda de datos y limitaciones en la generación de reportes.

Para abordar esta problemática, Matos (2022) planteó el desarrollo de un sistema web que permitiera automatizar los procesos administrativos relacionados con la gestión de expedientes. La investigación se desarrolló bajo un enfoque cuantitativo, de tipo aplicada y nivel descriptivo, con un diseño experimental que permitió evaluar los efectos de la implementación del sistema. Asimismo, se identificaron los requerimientos funcionales del sistema a partir de las necesidades de los usuarios, lo que permitió delimitar las funcionalidades clave del software.

En cuanto a la metodología, se utilizó Extreme Programming (XP), una metodología ágil que permitió desarrollar el sistema mediante iteraciones cortas y con constante retroalimentación del usuario. Durante la implementación, se emplearon tecnologías web como PHP, HTML, CSS, JavaScript y MySQL, así como el framework Laravel, lo que permitió construir un sistema robusto y funcional. El sistema desarrollado incluyó módulos para el registro y seguimiento de expedientes, control de diligencias, gestión de clientes, programación de citas y generación de reportes, permitiendo centralizar la información y mejorar su accesibilidad.

Asimismo, se realizaron pruebas unitarias y de aceptación con el fin de validar el correcto funcionamiento del sistema y verificar el cumplimiento de los requerimientos establecidos. También

se aplicaron encuestas de satisfacción a los usuarios finales, lo que permitió evaluar el impacto del sistema en los procesos del estudio jurídico.

Los resultados obtenidos (ver Tabla 6) evidenciaron mejoras significativas en la gestión de expedientes, logrando cubrir el 100% de los requerimientos planteados. Además, se registraron mejoras del 81.95% en el tiempo de registro de expedientes, 95.07% en la generación de reportes y 98.13% en la búsqueda de información, lo que demostró la eficiencia del sistema implementado.

Tabla 6

Rendimiento del enfoque propuesto en el conjunto de datos judiciales

Actividad	Tiempo Promedio Antes del Sistema (min)	Tiempo Promedio con el Sistema (min)	Mejora del tiempo de procesamiento de información
Registrar Expediente	13.3	2.4	81.95%
Reporte de Expediente	20.3	1	95.07%
Búsqueda	48.6	0.91	98.13%

Nota. Se muestran las métricas de tiempo obtenidas antes y después de la implementación de GESTEXP. Basado en: Matos, 2022

3.3.10. Sistema web para procuraduría pública en Perú

Este caso de éxito se centra en la investigación realizada por Walter Alejandro Castañeda Torres para la Universidad César Vallejo, titulada "Sistema de Información Web, Utilizando Metodología Scrum para la Gestión de Expedientes Judiciales en la Procuraduría Pública Regional de San Martín". El estudio aborda la transformación digital de una entidad gubernamental peruana responsable de la defensa jurídica del Estado en la región de San Martín. La investigación propone que la adaptación y el desarrollo de tecnologías de la información en instituciones públicas no solo automatizan procesos, sino que eliminan errores humanos derivados de la gestión manual, permitiendo que la información esté disponible en tiempo y forma para una ejecución coordinada de actividades legales.

La problemática identificada en la Procuraduría Pública Regional de San Martín radicaba en una gestión de expedientes sumamente tradicional y deficiente, basada en el uso de herramientas aisladas como formatos Excel, archivos físicos y cuadernos de cargo manuales para el registro de notificaciones. Esta falta de optimización generaba un manejo inadecuado de las notificaciones judiciales y una demora excesiva en la entrega de escritos, lo que frecuentemente provocaba el vencimiento de plazos legales. Las consecuencias eran críticas: pérdida de casos por rebeldía, archivamiento de procesos y perjuicios económicos significativos para el Estado debido al pago de reparaciones civiles o sentencias duplicadas por errores de registro. Además, el personal, incluyendo abogados y asistentes, manifestaba una alta insatisfacción debido a la opacidad y lentitud para consultar información esencial en entidades como RENIEC o bases de antecedentes judiciales.

Para dar solución a estos desafíos, la metodología aplicada fue de tipo cuantitativa y aplicada, utilizando un diseño de investigación experimental de nivel preexperimental con mediciones de "antes" y "después". El núcleo del desarrollo fue la metodología ágil Scrum, seleccionada por su capacidad para maximizar resultados mediante el trabajo en equipo y la gestión eficiente de requerimientos. El proceso incluyó seis etapas que fueron desde el diagnóstico inicial y la elaboración de requisitos funcionales hasta el desarrollo de prototipos, pruebas de funcionamiento y la entrega final del software.

La implementación tecnológica se materializó en un sistema de información web desarrollado con el lenguaje de programación Java y el gestor de base de datos PostgreSQL, elegido por su capacidad para ejecutar consultas complejas y su licencia de código abierto. El sistema fue estructurado en tres sprints principales: el primero enfocado en la autenticación y mantenimiento de perfiles; el segundo destinado al registro de expedientes, partes procesales y gestión de notificaciones; y el tercero orientado a la generación de reportes consolidados y seguimiento de procesos judiciales. La arquitectura contó con servidores de almacenamiento y respaldo para garantizar la integridad y seguridad de los datos sensibles. Además, se incluyeron servicios integrados (web services) para agilizar la búsqueda de datos de identidad y antecedentes, reduciendo drásticamente la carga operativa de los asistentes legales.

Los resultados obtenidos tras la implementación demostraron una mejora significativa en todos los indicadores evaluados. En cuanto a la eficiencia operativa, el tiempo promedio de consulta

a servicios integrados (RENIEC y antecedentes) se redujo de una media de 2 horas con 57 minutos a tan solo 58 segundos. La capacidad de respuesta institucional también mejoró: el porcentaje de demandas atendidas se elevó del 62.79% al 92.79%, mientras que las apelaciones generadas a tiempo subieron del 67.43% al 86.22%. En el ámbito de la satisfacción del usuario, los niveles de "Muy de acuerdo" crecieron significativamente en dimensiones críticas como la adaptabilidad (56.7%) y la usabilidad (36.7%), validando que el sistema es intuitivo y se ajusta fielmente a la dinámica de trabajo de la Procuraduría.

Tabla 7

Indicadores obtenidos para el sistema de gestión de la Procuraduría de San Martín

Indicador	Tipo Indicador	Pretest	Postest
Denuncias atendidas	Porcentaje	62.79	92.72
Apelaciones Generadas	Porcentaje	67.43	86.22
Encuesta de satisfacción (Usabilidad)	Porcentaje de respuestas en "Muy de acuerdo"	0	36.7
Encuesta de satisfacción (Adaptabilidad)	Porcentaje de respuestas en "Muy de acuerdo"	0	56.7
Encuesta de satisfacción (Operatividad)	Porcentaje de respuestas en "Muy de acuerdo"		30.0
Encuesta de satisfacción (Autenticidad)	Porcentaje de respuestas en "Muy de acuerdo"		33.3

Nota. Basado en: Castañeda, 2022

Capítulo IV

Metodología

4.1. Diseño de Investigación

4.1.1. Tipo de Investigación

La presente investigación es de tipo aplicada, ya que busca resolver un problema práctico mediante la reingeniería del sistema SIE. Este sistema incorpora herramientas tecnológicas innovadoras, como OCR y LLM, con el fin de mejorar la eficiencia, eficacia y rapidez en la Procuraduría. Según Hernández-Sampieri et al. (2014), la investigación aplicada se orienta a comprender y enfrentar problemas concretos en contextos reales. Asimismo, contribuye al diseño, implementación y mejora de programas institucionales. Este tipo de investigación integra diversas disciplinas y requiere la participación de los actores involucrados.

4.1.2. Nivel de Investigación

La investigación se desarrolla a nivel explicativo, ya que busca analizar la relación causal entre la reingeniería del sistema SIE y los cambios en la gestión de expedientes en la Procuraduría. Su objetivo es comprender cómo y por qué las modificaciones en el sistema influyen en los procesos institucionales. Según Hernández-Sampieri et al. (2014), este nivel de investigación permite explicar las causas y efectos de los fenómenos estudiados. Asimismo, contribuye a identificar las condiciones en las que se producen dichos cambios. De esta manera, se fortalece la comprensión de la relación entre ambas variables.

4.1.3. Enfoque de Investigación

La investigación adopta un enfoque cuantitativo, ya que permite analizar la relación entre variables medibles mediante datos numéricos y métodos estadísticos. Las variables son evaluadas a través de instrumentos de medición, lo que facilita la obtención de información objetiva. Según Hernández-Sampieri et al. (2014), este enfoque busca probar teorías de forma deductiva y controlar posibles sesgos. Asimismo, permite generalizar y replicar los resultados. De esta manera, se garantiza un análisis riguroso y confiable.

4.1.4. Diseño de Investigación

Se ha optado por un diseño cuasiexperimental de tipo mixto dado que no es posible realizar una asignación aleatoria de los sujetos de estudio.

El carácter mixto del diseño responde a la existencia de dos unidades de análisis diferenciadas: *abogados* y *expedientes*, para las cuales se emplearon esquemas de medición distintos, de acuerdo con la naturaleza de cada una. En el caso de la unidad de análisis *abogados*, se utilizó un diseño cuasiexperimental pretest–postest sin grupo de control, dado que fue posible medir las variables antes y después de la implementación del sistema, comparando los resultados entre ambos.

La notación para este diseño de investigación sería la siguiente:

$$E_n: O_1 \rightarrow x \rightarrow O_2$$

- E: Grupo experimental
- n: tamaño de la muestra de abogados
- X: Manipulación de la variable *Reingeniería del sistema web SIE*
- O_1 : Prueba realizada antes del experimento (Pretest)
- O_2 : Prueba realizada después del experimento (Postest)

Por otro lado, para la unidad de análisis expedientes, se aplicó un diseño cuasi-experimental de un solo test con grupo de control, debido a que, por la naturaleza del proceso, los expedientes son gestionados íntegramente bajo una única modalidad desde su ingreso, lo que imposibilita la aplicación de una medición previa sin alterar el flujo operativo normal. En este caso, el grupo de control estuvo conformado por los expedientes recepcionados en la Procuraduría durante febrero de 2025, los cuales fueron gestionados sin el sistema (periodo previo a la implementación). Por su parte, el grupo experimental estuvo conformado por los expedientes recepcionados durante marzo de 2025, los cuales fueron gestionados utilizando el sistema implementado (periodo posterior a la implementación).

Su notación es la siguiente:

$$C_{m/2}: O$$

$$E_{m/2}: x \rightarrow O$$

- C: Grupo de control. Conformado por el 50% de la muestra.
- E: Grupo experimental. Conformado por el 50% de la muestra.
- m: tamaño de la muestra de expedientes.
- X: Manipulación de la variable *Reingeniería del sistema web SIE*
- O : Prueba realizada (Test)

4.2. Población y Muestra

4.2.1. Población

- Población 1: Abogados que ejercen su labor en la actualidad en el área penal de la Procuraduría. Actualmente el valor asciende a 38 trabajadores.
- Población 2: Expedientes judiciales del área penal de la Procuraduría Pública del Sector Interior – MININTER gestionados mensualmente durante el año 2024. El valor asciende a 1949 expedientes.

4.2.2. Muestra

- Muestra 1: No aplica. Se realizará un estudio censal.
- Muestra 2: Dado que la población es de tipo finita, se hará uso de la siguiente fórmula estadística:

$$n = \frac{Z^2 pq N}{E^2 (N - 1) + Z^2 pq}$$

- Nivel de confianza: 99%
- Valor de Z: 2.576
- Proporción poblacional que tiene la característica de interés p: 0.5
- Proporción poblacional que no tiene la característica de interés p: 0.5
- Tolerancia de error: $0.05 = 5\%$
- Tamaño de la población N: 1949
- Tamaño de la muestra: n

$$n = \frac{(2.576)^2(0.5)(0.5)(1049)}{(0.05)^2(1948) + (2.576)^2(0.5)(0.5)}$$

$$n = 495.2227889 \approx 496$$

A partir de lo anterior, podemos visualizar que el tamaño de la muestra a utilizar en la experimentación de este proyecto es de 496 expedientes del área penal.

4.2.3. Unidad de Análisis

- Unidad 1: Abogados del área penal de la Procuraduría Pública del Sector Interior – MININTER
- Unidad 2: Expedientes judiciales del área penal de la Procuraduría Pública del Sector Interior – MININTER

4.2.4. Diseño Muestral

En la presente investigación se consideran dos poblaciones de estudio. La primera está conformada por los abogados del área penal, mientras que la segunda corresponde a los expedientes judiciales gestionados por la Procuraduría. En función de las características de cada población, se aplicaron distintos procedimientos de selección muestral.

Para la primera población, conformada por los abogados del área penal, se optó por aplicar un muestreo censal, debido a que el número total de integrantes es reducido (38 abogados). En este caso, se consideró pertinente incluir a la totalidad de la población, con el propósito de obtener información completa y evitar posibles sesgos derivados de una selección parcial.

Por otra parte, para la segunda población, conformada por los expedientes judiciales, se empleó un muestreo probabilístico de tipo aleatorio simple, con la finalidad de obtener una muestra representativa. Este método garantiza que todos los expedientes tengan la misma probabilidad de ser seleccionados, permitiendo realizar inferencias válidas sobre la población sin necesidad de analizar la totalidad de los casos.

Según Hernández-Sampieri et al. (2014), en las investigaciones cuantitativas orientadas a describir o analizar relaciones entre variables, el uso de muestras probabilísticas resulta fundamental, ya que permite obtener estimaciones precisas de los parámetros poblacionales y

reducir el error muestral. Asimismo, este enfoque contribuye a fortalecer la validez de los resultados obtenidos.

Asimismo, para la identificación de los elementos de ambas poblaciones, se utilizó la base de datos institucional de la Procuraduría como marco muestral. Dicho registro permitió delimitar adecuadamente los sujetos de estudio y garantizar el cumplimiento de los criterios de inclusión establecidos. Para la población 2 se consideró únicamente los expedientes judiciales del área penal en proceso de trámite hasta el mes de marzo del 2025, que cuenten con información completa, pertenezcan al proceso común, se encuentren registrados en el sistema institucional y correspondan a alguno de los veinte (18) delitos principales atendidos por la Procuraduría, tales como la falsificación o adulteración de documentos militar-policial, la conducción en estado de ebriedad, la tenencia ilegal de armas de fuego y/o municiones y el tráfico ilícito de migrantes. La relación completa de los delitos considerados se presenta en el Anexo 1. Por otro lado, para la población 1, se consideró a los abogados activos del área penal que participan directamente en la gestión de dichos expedientes.

4.3. Matriz de consistencia

Tabla 8

Matriz de consistencia

Problemas	Objetivos	Hipótesis	Variables	Metodologías
General			VI:	Tipo de Investigación
¿De qué manera la reingeniería del sistema web SIE permite mejorar la gestión de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER?	Realizar una reingeniería del sistema web SIE para mejorar la gestión de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.	La reingeniería del sistema web SIE mejora la gestión de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.	Reingeniería del sistema web SIE	El estudio será de tipo cuasiexperimental.
Específicos			VD:	Población y Muestra
¿De qué manera la reingeniería del sistema web SIE permite mejorar la recepción y consulta de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER?	Realizar una reingeniería del sistema web SIE para mejorar la recepción y consulta de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.	La reingeniería del sistema web SIE mejora la recepción y consulta de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.	Gestión de expedientes judiciales	Población N°1: 38 abogados del área penal de la Procuraduría Pública del Sector Interior – MININTER Muestra N°1: Censal
¿De qué manera la reingeniería del sistema web SIE permite mejorar la elaboración y envío de recursos legales de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER?	Realizar una reingeniería del sistema web SIE para mejorar la elaboración y envío de recursos legales de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.	La reingeniería del sistema web SIE mejora la elaboración y envío de recursos legales de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.		Población N°2: 1949 expedientes judiciales del área penal de la Procuraduría Pública del Sector Interior – MININTER Muestra N°2: 496 expedientes judiciales del área penal de la

¿De qué manera la reingeniería del sistema web SIE permite mejorar la planificación y participación en audiencias de los miembros de la Procuraduría Pública del Sector Interior - MININTER?	Realizar una reingeniería del sistema web SIE para mejorar la planificación y participación en audiencias de los miembros de la Procuraduría Pública del Sector Interior - MININTER.	La reingeniería del sistema web SIE mejora la planificación y participación en audiencias de los miembros de la Procuraduría Pública del Sector Interior - MININTER.
¿De qué manera la reingeniería del sistema web SIE permite mejorar la revisión y seguimiento de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER?	Realizar una reingeniería del sistema web SIE para mejorar la revisión y seguimiento de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.	La reingeniería del sistema web SIE mejora la revisión y seguimiento de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.

Nota. Elaboración Propia.

4.4. Operacionalización de variables

Tabla 9

Definición operacional de la variable independiente

Variable	Definición operacional	Indicador	Técnica	Escala
Reingeniería del sistema web SIE	Transformación del sistema SIE anterior en un sistema modernizado, con el propósito de adaptarlo a las nuevas necesidades tecnológicas y organizacionales de los usuarios	Estado del sistema	Análisis documental	0 = Antes / 1 = Después

Nota. Elaboración Propia.

Tabla 10

Definición operacional de la variable dependiente

Variable	Definición Operacional	Dimensión	Indicadores	Instrumento	Tipo	Escala de Medición	Categoría	Anexo
<i>Gestión de expedientes judiciales</i>	Conjunto de procesos y actividades destinadas a administrar, organizar y controlar la información	Recepción y consulta de expedientes judiciales	Nivel de percepción sobre la completitud, precisión y accesibilidad de la documentación de los expedientes recibidos.	Cuestionario	Cuantitativo	Razón	Numérico	Anexo 2 (Item 1,2,3,4)

relacionada con casos judiciales a lo largo de su ciclo de vida.		Tiempo promedio de recepción de expedientes.	Cuestionario	Cuantitativo	Razón	Numérico	Anexo 3
			Ficha de registro de datos	Cuantitativo	Razón	Numérico	
	Elaboración y envío de recursos legales	Nivel de percepción sobre la rapidez en la elaboración de recursos legales	Cuestionario	Cuantitativo	Razón	Numérico	Anexo 2 (Item 5,6)
		Tiempo promedio de elaboración de recursos legales.	Ficha de registro de datos	Cuantitativo	Razón	Numérico	Anexo 3
		Nivel de aceptación de recursos legales por otras entidades	Cuestionario	Cuantitativo	Razón	Numérico	Anexo 2 (Item 7,8)
	Planificación y participación en audiencias	Nivel de percepción sobre el cumplimiento de los plazos en la programación de audiencias.	Cuestionario	Cuantitativo	Razón	Numérico	Anexo 2 (Item 9)
		Nivel de asistencia a audiencias de los	Cuestionario	Cuantitativo	Razón	Numérico	Anexo 2 (Item 10, 11,12)

		procuradores adjuntos.					
	Revisión y seguimiento de expedientes	Grado de identificación de tendencias o patrones recurrentes en los expedientes.	Cuestionario	Cuantitativo	Razón	Numérico	Anexo 2 (Item 13)
		Grado de seguimiento del estado de los expedientes	Cuestionario	Cuantitativo	Razón	Numérico	Anexo 2 (Item 14, 15,16)

Nota. Elaboración Propia.

4.5. Técnicas e Instrumentos de Recolección de Datos

4.5.1. Técnicas

- **Encuestas:** Se utilizó la técnica de la encuesta para recopilar información de los abogados del área penal (Población 1) que utilizan el sistema web SIE en la Procuraduría Pública del Sector Interior. Esta técnica permitirá obtener datos sobre sus percepciones, nivel de satisfacción, experiencia de uso y valoración del sistema. Según Hernández-Sampieri et al. (2014), las encuestas permiten recopilar información sobre actitudes, opiniones y características de una población determinada, facilitando la obtención de datos cuantitativos susceptibles de análisis estadístico.
- **Análisis documental:** Se empleó la técnica del análisis documental para el estudio de los expedientes judiciales (Población 2), con el fin de examinar los tiempos operativos y otros datos relevantes registrados en el sistema. Esta técnica permitirá recolectar información objetiva a partir de logs del sistema.

4.5.2. Instrumentos

- **Cuestionario Estructurado:** Se diseñó un cuestionario estructurado dirigido a los abogados del área penal, orientado a evaluar su percepción sobre la recepción de expedientes, la elaboración de recursos legales, la participación en audiencias y el seguimiento de los expedientes. El instrumento consta de 16 ítems, distribuidos en cuatro dimensiones, con cuatro preguntas por cada una. Las preguntas fueron formuladas bajo una escala tipo Likert de cinco niveles, que van desde “Nunca” (1) hasta “Siempre” (5). (Ver anexo 2)
- **Ficha de registro de datos:** Se utilizó una ficha de registro de datos para poder obtener los siguientes indicadores de tiempo:

I1: Tiempo de recepción de documentos

I2: Tiempo promedio de elaboración de recursos legales

El indicador I1 mide el tiempo promedio que le toma a Mesa de partes recepcionar un documento relacionado a un expediente, mientras que I2 mide el tiempo promedio que le toma a un abogado elaborar un recurso legal asociado a un expediente.

Ambos indicadores se midieron tanto en el grupo de control como en el grupo experimental. Sin embargo, cabe mencionar que en el grupo experimental el tiempo considerado incluye tanto el tiempo que le toma al sistema procesar o generar el documento más el tiempo que le toma al usuario revisar o modificar la información.

Por lo que la notación final de ambos indicadores sería la siguiente:

$$I = I_{\text{automatizacion}} + I_{\text{manual}}$$

donde $I_{\text{automatizacion}}$ es 0 en el grupo de control.

4.6. Técnicas para el Procesamiento de la Información

Se emplearon diversas estrategias de procesamiento de datos con el objetivo de analizar los datos recopilados y obtener conclusiones significativas. Para llevar a cabo este proceso, se utilizará el software estadístico SPSS, reconocido por su facultad de realizar evaluaciones estadísticas de nivel avanzado y generar visualizaciones claras y precisas. Las técnicas de procesamiento de la información seleccionadas para este proyecto son las siguientes:

- **Análisis Descriptivo:** El análisis descriptivo será utilizado para explorar y resumir las características principales a partir de la información recogida por encuestas y observación. Mediante el uso de SPSS, se calcularán herramientas estadísticas como promedio, mediana, moda, variabilidad estándar y el rango intercuartílico, que proporcionarán una descripción detallada de la relativa a la forma en que los datos se distribuyen y su grado de variación.
- **Análisis Inferencial:** Con la finalidad de determinar cómo la reingeniería del sistema SIE web repercute en la gestión de expedientes judiciales, se utilizará un análisis estadístico inferencial. Este análisis permitirá determinar si existen diferencias significativas en el manejo de expedientes judiciales antes y tras la puesta en marcha del sistema mejorado, lo que indicaría el efecto de la reingeniería aplicada en dicha gestión.

Capítulo V

Implementación

5.1. Ingeniería Inversa (AS - IS)

5.1.1. Conceptualización y Requerimientos del sistema anterior

El sistema anterior satisface 4 necesidades que se indican en la siguiente tabla:

Tabla 11

Necesidades satisfechas en el sistema anterior

Necesidad	Descripción
Visualización y búsqueda de legajos	Los usuarios necesitan consultar legajos mediante tablas, filtros y opciones de búsqueda avanzada.
Gestión documental del legajo	Los usuarios necesitan registrar, editar y eliminar documentos de entrada y salida asociados a un legajo.
Control del estado y exportación	Los usuarios necesitan actualizar el estado del legajo y la exportación de su información en formatos digitales.
Creación y actualización de legajos	Los usuarios necesitan registrar nuevos legajos e incorporar documentos en legajos existentes.

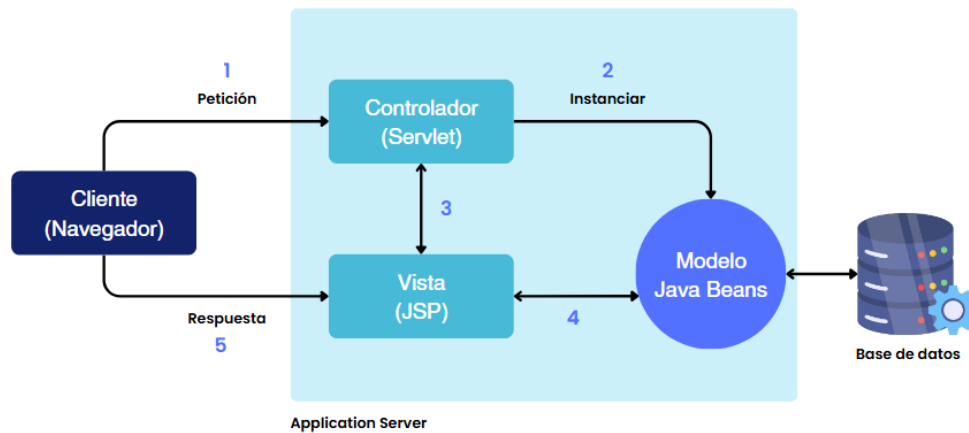
Nota. Elaboración propia.

5.1.2. Diseño del sistema anterior

El sistema anterior fue desarrollado sobre una arquitectura web tradicional basada en Java, siguiendo el patrón Modelo–Vista–Controlador (MVC). Las vistas se renderizan en el servidor mediante JavaServer Pages, mientras que los controladores gestionan la lógica de negocio y el flujo de la aplicación. Este diseño sigue un framework Java antiguo llamado Apache Struts, ampliamente utilizado en sistemas institucionales y gubernamentales. Asimismo, el sistema SIE sigue una arquitectura monolítica, con generación de contenido del lado del servidor y una separación entre presentación y procesamiento, propia de aplicaciones desarrolladas en etapas previas a las arquitecturas modernas basadas en APIs REST y frontends desacoplados.

Figura 17

Arquitectura del sistema anterior



Nota. Elaboración propia.

5.1.3. Funcionalidades del sistema anterior

El sistema SIE cuenta con una pantalla principal que presenta dos módulos: Legajos del área penal y Consolidado de información/Reportes. El módulo de Legajos del área penal constituye el núcleo operativo del sistema, mientras que el módulo de Consolidado brinda una función de reportes institucionales.

Módulo Recepción y Consulta de Legajos. Se puede acceder al módulo de recepción y consulta a través del menú “Penal”. Al acceder a esta opción, el sistema muestra una interfaz general que permite realizar búsquedas mediante diversos criterios, tales como número de legajo, fecha de registro, tipo de proceso y tipo de documento. Como resultado, se presenta una tabla consolidada que contiene todos los legajos registrados dentro del ámbito penal.

Cada registro permite realizar operaciones básicas de administración, entre las cuales se encuentran la modificación del estado del expediente (por ejemplo, en curso, archivado o finalizado) y la visualización de los documentos asociados.

Figura 18

Pantalla Legajos del área penal

Nro. Legajo	Entidad	Región	Dependencia	DetEnt	Dist. Judicial	TipDoc	Tipo	Proceso	Motivo de Ingreso	Carpeta	Nro. Exp	Partes	Estado	FecArc	Grado Imp	Ols	ResPrincipal
P-2024009789	PODER JUDICIAL	AREQUIPA			2° JUZGADO DE LA INVESTIGACION PREPARATORIA DE PAUCARPATA	RESOLUCION	CASO PENAL	PROCESO PENAL N.C.P.P.	AUDIENCIA DE CONTROL DE ACUSACION				EN TRAMITE		NORMAL	SINOE	
P-2024009788	MINISTERIO PUBLICO	LIMA			PRIMERA FISCALIA CORPORATIVA PENAL DE MIRAFLORES	DISPOSICION FISCAL	CASO PENAL	PROCESO PENAL N.C.P.P.	AMPLIACION DE CARPETA FISCAL				EN TRAMITE		NORMAL	CONTRA "L.Q.R.L."	

Nota. Captura de pantalla de legajos del área penal del sistema SIE. Acceso institucional, 2025

Figura 19

Pantalla Documentos de Ingreso

Código	Número	Dependencia Legajo	Dist. Judicial Legajo	Dist. Judicial Expediente	Tipo de Remitente	Remitente	Situación Jurídica	Domicilio Procesal	Grado	Funcionario	Tipo de Documento	Nro. de Documento	Fecha de Documento	Sumilla	Plazo HT	Monto de la reparación civil	Número y/o de depósito judicial	Observación	Registrador	Fec Res	
	900556	MISERACIONES - SUPERINTENDENCIA NACIONAL DE MISERACIONES									RUD	2024-0002964	15/01/2024	INICIO DE ACCIONES		9,0					15/01/2024 09:4

Nota. Captura de pantalla de los documentos de ingreso de un expediente en el sistema SIE. Acceso institucional, 2025

Figura 20

Pantalla Documentos de Salida

Código	Tipo Destinatario	Destinatario	Tipo Documento	Tipo Escrito	Fecha del Documento	Nro.	Sumilla	Folios	Abogado Redactor	Archivo Adjunto	Registrador	Fecha Registro	Otros Anexos	Modificar
257612	MININTER/PPH	MIGRACIONES - SUPERINTENDENCIA NACIONAL DE MIGRACIONES	FORMATO DE ARCHIVO-PSI		20/06/2024	-2024 IN/PSI	ARCHIVO POR PRESCRIPCION	03		Descargar		20/06/2024 15:09:07	Ver	☐
257506	MININTER/PPH	MIGRACIONES - SUPERINTENDENCIA NACIONAL DE MIGRACIONES	FORMATO DE ARCHIVO-PSI		19/06/2024	-2024 IN/PSI	SOLICITA ARCHIVO	3		Descargar		19/06/2024 22:37:57	Ver	☐

Nota. Captura de pantalla de los documentos de salida de un expediente en el sistema SIE. Acceso institucional, 2025

Dentro de este sistema, a cada expediente judicial se le asigna un código interno denominado “Número de Legajo”, el cual permite su identificación y seguimiento institucional. Inicialmente, el sistema contemplaba el registro detallado de aproximadamente veinte campos informativos por expediente además del “Número de legajo”, incluyendo datos procesales, administrativos y de identificación. Sin embargo, debido a la alta carga operativa y al tiempo requerido para completar dichos registros, este proceso se ha visto progresivamente simplificado, limitándose en muchos casos a la carga directa de los documentos, sin una estructuración exhaustiva de la información.

Módulo de Reportes de legajos. Las opciones “Cons.” y “Resp.” permiten a los usuarios acceder a diversas opciones de exportación y generación de reportes. Entre los principales reportes disponibles se encuentran:

- Consolidado de documentos de ingreso o documentos de entrada.
- Consolidado principal de legajos.
- Consolidado por rangos de fecha y hora.
- Consultas por registrador y usuario responsable, entre otros.

Figura 21

Pantalla Reportes

Nota. Captura de pantalla del apartado de reportes. Acceso institucional, 2025

5.2. Alteración (Transformación)

5.2.1. Análisis de Problemas y Limitaciones

En base a reuniones realizadas con los usuarios se identificaron 9 necesidades actuales, de las cuales solo 4 eran cubiertas por funcionalidades del sistema anterior. El detalle se muestra a continuación:

Tabla 12

Análisis de necesidades satisfechas en el sistema anterior

Necesidad	Descripción	¿Satisfecha en el sistema actual?	Observación
Visualización y búsqueda de legajos	Los usuarios necesitan consultar legajos mediante tablas, filtros y opciones de búsqueda avanzada.	Sí	-
Gestión documental del legajo	Los usuarios necesitan registrar, editar y eliminar documentos de entrada y salida asociados a un legajo.	Sí	Los documentos de salida no se generan automáticamente sino que se suben manualmente al sistema.

Control del estado y exportación	Los usuarios necesitan actualizar el estado del legajo y la exportación de su información en formatos digitales.	Sí	-
Creación y actualización de legajos	Los usuarios necesitan registrar nuevos legajos e incorporar documentos en legajos existentes.	Sí	-
Extracción rápida de información contenida en los documentos	Los usuarios necesitan registrar en el sistema, de manera rápida, la información proveniente de documentos de entrada	No	El llenado de campos se realiza de manera completamente manual por lo que demora un tiempo considerable
Validación y corrección de datos	Los usuarios necesitan revisar, corregir y confirmar la información extraída rápidamente.	No	La actualización de los campos de un legajo se realiza de manera manual
Programación de audiencias	Los usuarios necesitan programar audiencias relacionadas a los legajos	No	El sistema anterior no permite programar o registrar audiencias relacionadas a un legajo.
Modificación y reprogramación	Los usuarios necesitan actualizar, reasignar fechas o horarios y eliminar audiencias.	No	El sistema anterior no permite actualizar las audiencias relacionadas a un legajo.
Visualización de indicadores	Los usuarios necesitan ver métricas clave de la gestión de legajos	No	El sistema anterior no cuenta con métricas clave.

Nota. Elaboración propia.

5.2.2. Propuesta de Mejora

La Tabla 13 presenta un resumen de las propuestas de mejora para el nuevo sistema, las cuales están orientadas a fortalecer sus capacidades operativas. En primer lugar, se plantea que el nuevo sistema incorpore nuevas funcionalidades en el módulo de recepción de legajos, mediante la implementación de un servicio backend basado en modelos de lenguaje (LLM), que permitirá automatizar el proceso de recepción y registro de expedientes con un enfoque de validación

humana (human-in-the-loop). Asimismo, el módulo de consulta de legajos será potenciado con un servicio para la generación automática de recursos legales y su posterior envío a través del sistema.

Por otro lado, se propone la creación de un nuevo módulo para la gestión de audiencias, el cual integrará un calendario institucional que facilitará la planificación, edición y control de las participaciones procesales. Finalmente, el módulo de reporte de legajos será reforzado mediante la implementación de un dashboard de métricas.

Tabla 13

Mejoras propuestas para el nuevo sistema

Necesidad	Módulo en el nuevo sistema	Descripción de la propuesta
Visualización y búsqueda de legajos	Consulta Legajo	Se mantienen las funcionalidades.
		Nueva funcionalidad que permita la generación automática de recursos legales.
Gestión documental del legajo	Consulta Legajo	Nueva funcionalidad que permita el envío de recursos legales a través del sistema.
Control del estado y exportación	Consulta Legajo	Se mantienen las funcionalidades.
Creación y actualización de legajos	Recepción Legajo	Se mantienen las funcionalidades.
Extracción rápida de información contenida en los documentos	Recepción Legajo	Nueva funcionalidad para la identificación automática de entidades utilizando OCR y un modelo LLM.
Validación y corrección de datos	Recepción Legajo	Nueva funcionalidad para observar la información extraída automáticamente y corregir en caso sea necesario.
Programación de audiencias	Audiencias	Nuevas funcionalidades para crear audiencias en formato de calendar.
Modificación y reprogramación	Audiencias	Nuevas funcionalidades para modificar y eliminar audiencias.
Visualización de indicadores	Reportes	Dashboard con métricas clave de la gestión de legajos

Nota. Elaboración propia.

5.3. Ingeniería Directa (TO - BE)

5.3.1. Conceptualización y Requerimientos del nuevo sistema

5.3.1.1. Definición de user stories. El desarrollo del sistema se inició con la identificación de las necesidades del usuario, las cuales fueron descompuestas posteriormente en user stories, siguiendo la metodología SCRUM.

Tabla 14

Necesidades y user stories en el nuevo sistema

Código	Necesidad	Descripción	User Stories Asociadas	Módulo
N01	Visualización y búsqueda de legajos	Los usuarios necesitan consultar legajos mediante tablas, filtros y opciones de búsqueda avanzada.	SIE-11 a SIE-18	Consulta Legajo
N02	Gestión documental del legajo	Los usuarios necesitan registrar, editar y eliminar documentos de entrada y salida asociados a un legajo.	SIE-19 a SIE-25	Consulta Legajo
N03	Control del estado y exportación	Los usuarios necesitan actualizar el estado del legajo y la exportación de su información en formatos digitales.	SIE-25 a SIE-29	Consulta Legajo
N04	Creación y actualización de legajos	Los usuarios necesitan registrar nuevos legajos e incorporar documentos en legajos existentes.	SIE-04 a SIE-10	Recepción Legajo
N05	Extracción rápida de información contenida en los documentos	Los usuarios necesitan registrar en el sistema, de manera rápida, la información proveniente de documentos de entrada	SIE-18 a SIE-21	Recepción Legajo
N06	Validación y corrección de datos	Los usuarios necesitan revisar, corregir y confirmar la información extraída rápidamente.	SIE-32 a SIE-38	Recepción Legajo

N07	Programación de audiencias	Los usuarios necesitan programar audiencias relacionadas a los legajos	SIE-41 a SIE-42	Audiencias
N08	Modificación y reprogramación	Los usuarios necesitan actualizar, reasignar fechas o horarios y eliminar audiencias.	SIE-43 a SIE-47	Audiencias
N09	Visualización de indicadores	Los usuarios necesitan ver métricas clave de la gestión de legajos	SIE-39 a SIE-41	Reportes

Nota. Elaboración propia.

En total, se definieron 47 user stories, las cuales fueron priorizadas mediante la técnica MoSCoW y estimadas en puntos de historia, permitiendo una planificación eficiente del trabajo. Para su gestión y seguimiento, se utilizó la herramienta Jira, facilitando el control del avance, la asignación de tareas y la coordinación del equipo.

Dentro de la herramienta, cada una de estas user stories contó con la siguiente estructura básica:

- Título de la user story
- Descripción o aspectos clave
- Estado (Tarea por hacer, en curso y finalizado)
- Tareas asociadas o Tareas secundarias
- Persona asignada a la user story
- Porcentaje de avance
- Story point estimate
- Clasificación MoSCoW

A continuación, se presentan algunos ejemplos de historias de usuario del proyecto.

Figura 22

User Story SIE-12

Espacio / SIE / SIE-50 / SIE-12

Como abogado, puedo filtrar los legajos para una búsqueda rápida

Descripción

Tanto el filtro para la Pantalla de **Mis Legajos** como para la pantalla de **Todos los legajos** debe contener los siguientes componentes:

- Search para **Legajo**
- Select para **Tipo de Caso**
- Search para **Número de caso**
- Range Picker para la **Fecha de Inicio** y **Fecha de Fin** de búsqueda

En el caso de la pantalla **Todos los legajos** se le debe añadir un Select para la selección del Abogado designado.

Subtareas

Añadir subtarea

Actividades vinculadas

Añadir actividad vinculada

Finalizado

✓ Listo

⚡

⌘ Mejorar el tipo de actividad Historia

Detalles

Persona asignada

N

nabia.pachas

Prioridad

= Medium

Principal

SIE-50 Consulta de legajos

Fecha de vencimiento

11 oct 2024

Etiquetas

MustHave

Team

Ninguno

Start date

10 oct 2024

Sprint

Sprint 3

Story point estimate

Ninguno

Informador

N

nabia.pachas

Nota. Elaboración Propia.

Figura 23

User Story SSIE-35

SSIE-3 / SSIE-35

Como abogado, quiero exportar el detalle de un legajo en PDF

Descripción

Para el diseño del PDF tomar en cuenta el formato adjunto en lineas más abajo

Start date

11 oct 2024

Fecha de vencimiento

12 oct 2024

Actividades secundarias

100 % completado

Actividad	Pri...	Pe...	Estado
SSIE-104 APIS de Carga de...	=	N	FINALIZADA
SSIE-103 API Download...	=	N	FINALIZADA

Finalizada

✓ Listo

⚡

Detalles

Persona asignada

N

Nabia Pachas

Etiquetas

MustHave

Principal

SSIE-3 Módulo 2- Consulta de Leg

Team

Añadir equipo

Sprint

SSKYP Sprint 3

Story point estimate

5

Nota. Elaboración Propia.

Figura 24

User Story SIE-7

Espacio / SIE / SIE-49 / SIE-7

Como mesa de partes, el sistema puede identificar los datos generales del documento

Descripción

En el API de Python se está identificando los datos Generales del legajo proporcionado. Para ello se necesita extraer la siguiente información:

- Lugar o distrito judicial
- Fiscalía (nombre)
- Fiscalía(nombreFiscal, tipoFiscal)
- Imputados (nombre completo, documento de identidad y sus delitos)
- Agraviados (nombre completo y documento de identidad)
- Hecho

Subtareas

Actividades vinculadas

Vincular actividades similares

Compartir 5 resultados

Finalizado

✓ Listo

⚡

⌘ Mejorar el tipo de actividad Historia

▼ Detalles

Persona asignada

kevinmiguel.ortiz

Asignarme a mí

Prioridad

Medium

Principal

SIE-49 Procesamiento de documentos

Fecha de vencimiento

14 oct 2024

Etiquetas

MustHave

Team

Ninguno

Start date

12 oct 2024

Sprint

Sprint 1

Story point estimate

Ninguno

Informador

nabia.pachas

Nota. Elaboración Propia.

5.3.1.2.Definición del product backlog. El product backlog está compuesto por las 47 user stories definidas en el apartado anterior.

Figura 25

Product Backlog del proyecto.

Backlog (53 actividades)				0	0	0
SIE-1	Como usuario, puedo ingresar en el sistema de acuerdo a mi pe...	LOGIN Y CONFIGURA...	POR HACER	1 oct		
SIE-2	Como usuario, puedo navegar en la plataforma a través de un menú	LOGIN Y CONFIGURA...	POR HACER	4 oct		
SIE-3	Como usuario, puedo observar mi información personal cuando esté logeado	LOGIN Y CONFIGURA...	POR HACER	7 oct		
SIE-4	Como mesa de partes, mis legajos son procesados con OCR	PROCESAMIENTO DE ...	POR HACER	1 oct		
SIE-5	Como mesa de partes, el sistema puede identificar el remitente que envía los documentos	PROCESAMIENTO DE ...	POR HACER	6 oct		
SIE-6	Como mesa de partes, el sistema puede identificar el tipo de documento y su categoría o clase	PROCESAMIENTO DE ...	POR HACER	9 oct		
SIE-7	Como mesa de partes, el sistema puede identificar los datos generales del documento	PROCESAMIENTO DE ...	POR HACER	14 oct		
SIE-8	Como mesa de partes, el sistema puede identificar el estado del caso judicial	PROCESAMIENTO DE ...	POR HACER	23 oct		
SIE-9	Como mesa de partes, el sistema puede identificar las audiencias notificadas	PROCESAMIENTO DE ...	POR HACER	30 oct		
SIE-10	Como usuario, quiero poder asignar un legajos a los abogados de manera secuencial	PROCESAMIENTO DE ...	POR HACER	7 nov		
SIE-11	Como abogado, puedo acceder tanto a mis legajos como al listado general de legajos	CONSULTA DE LEGAJ...	POR HACER	9 oct		
SIE-12	Como abogado, puedo filtrar los legajos para una búsqueda rápida	CONSULTA DE LEGAJ...	POR HACER	10 oct		

Nota. Elaboración Propia.

5.3.1.3. Definición del sprint backlog. Para la definición del sprint backlog, se consideró un total de 6 sprints, tal como se especifica a continuación:

Tabla 15

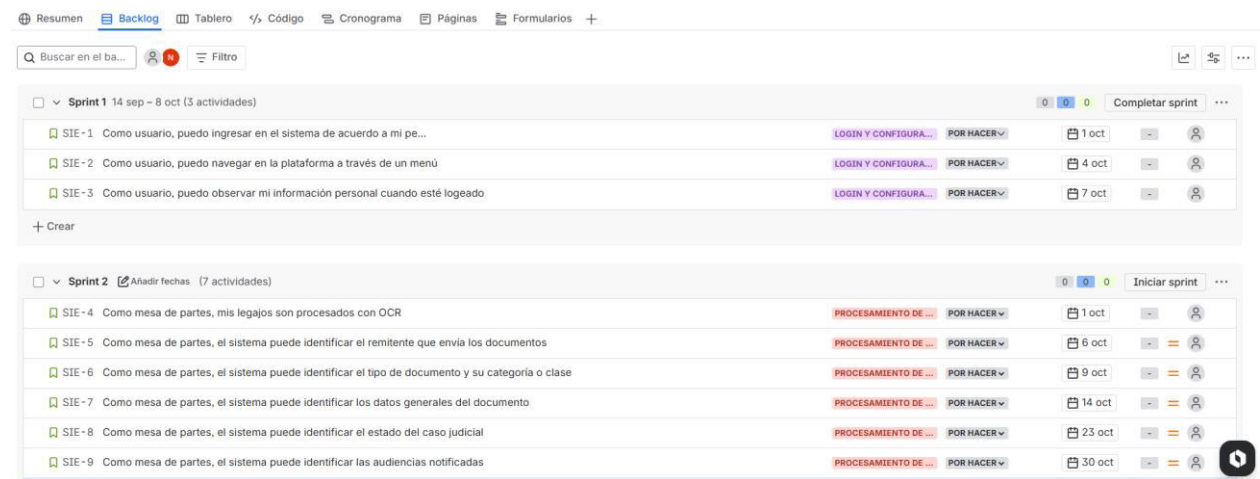
Sprint backlog del nuevo sistema

Código	Sprint	User stories asociadas	Stack
S1	Procesamiento de Documentos	SIE-01 a SIE-03	Python
S2	Login y configuraciones	SIE-04 a SIE-10	C# y React
S3	Consulta de Legajos	SIE-11 a SIE-29	C# y React
S4	Recepción de Legajos	SIE-30 a SIE-38	C# y React
S5	Reporte	SIE-39 a SIE-41	C# y React
S6	Audiencias	SIE-41 a SIE-47	C# y React

Nota. Elaboración Propia.

Figura 26

Sprint Backlogs

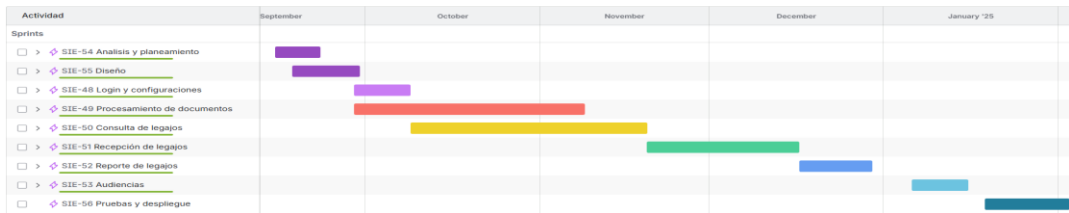


Nota. Elaboración Propia.

5.3.1.4.Cronograma. A partir de las user stories agrupadas por sprint Backlogs se definió el cronograma del proyecto. Este tiene una fecha de inicio de 14/09/2024 y una fecha de finalización de 11/02/2025. Cabe destacar que además del tiempo necesario para la implementación de las user stories, también se consideró el tiempo dedicado a operaciones previas como análisis y diseño; y a actividades posteriores como pruebas funcionales y despliegue.

Figura 27

Cronograma del Proyecto.



Nota. Elaboración Propia.

5.3.2. Diseño del nuevo sistema

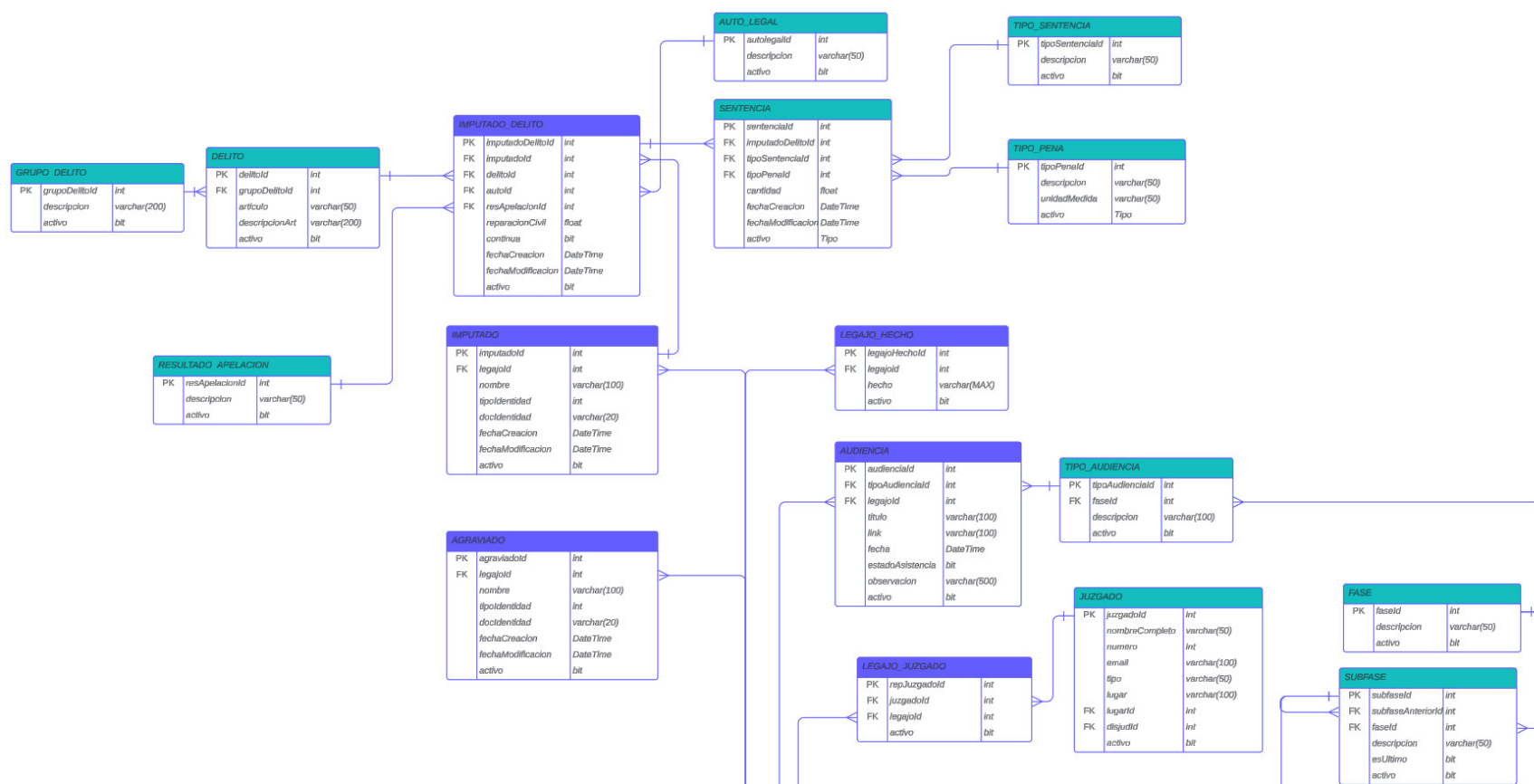
5.3.2.1.Modelo de datos. El modelo de datos para el sistema SIE incluye las principales entidades implicadas en la gestión de expedientes judiciales del sector interno del MININTER. Entre las más importantes se incluyen:

- Campos de Legajo: representa casos legales e incluye características como LegajoID (clave primaria), FechaRegistro, EstadoId, SubfaseId, entre otros. Cada legajo puede contener enlaces a uno o más documentos que puedan cubrirlo.
- Campos de Audiencia: vinculados a la tabla principal y que tienen atributos como AudienciaID (clave primaria), FechaAudiencia, TipoAudiencia y Observaciones. Además, está vinculado con la clave LegajoID.
- Campos de Usuario: Son las personas que administran los legajos ya sea como mesa de partes o abogados. Los atributos incluyen ID de usuario, nombre, función, correo electrónico, etc. Cada usuario tiene acceso a múltiples legajos. Pero cada legajo tiene un solo responsable.

El detalle completo del modelo de datos se puede apreciar en la Figura 28 y Figura 29.

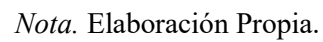
Figura 28

Modelo de Datos – Sección 1.



Nota. Elaboración Propia.

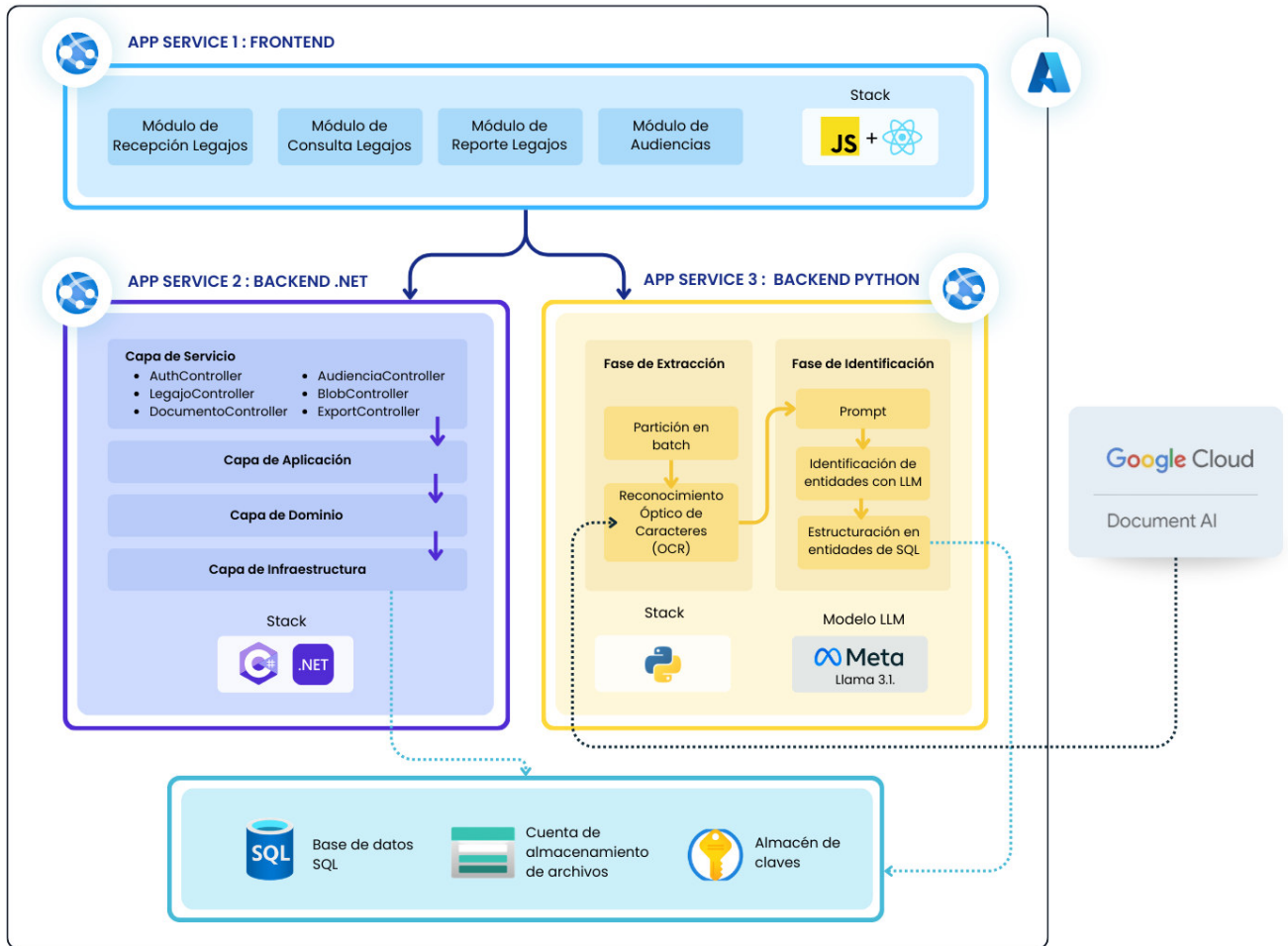
Modelo de Datos – Sección 2.



5.3.2.2.Arquitectura General.

Figura 30

Arquitectura de la solución.



Nota. Elaboración Propia.

La solución propuesta se implementa bajo una arquitectura distribuida basada en servicios en la nube, utilizando la plataforma Azure. El sistema se estructura en tres servicios principales: frontend, backend en .NET y backend en Python, los cuales interactúan entre sí mediante interfaces de comunicación definidas.

Capa de Presentación (App Service 1: Frontend). La capa de presentación corresponde al servicio frontend, desarrollado con JavaScript y React. Este componente permite la interacción directa con los usuarios mediante los módulos de recepción, consulta, reporte de legajos y gestión de audiencias.

Capa de Lógica de Negocio (App Service 2: Backend .NET). El backend principal está implementado en .NET y sigue una arquitectura en capas, conformada por las capas de servicio, aplicación, dominio e infraestructura. La capa de servicio expone los endpoints del sistema mediante controladores especializados, como AuthController, LegajoController, DocumentoController, AudienciaController, BlobController y ExportController. Estos controladores gestionan las solicitudes provenientes del frontend y coordinan su procesamiento.

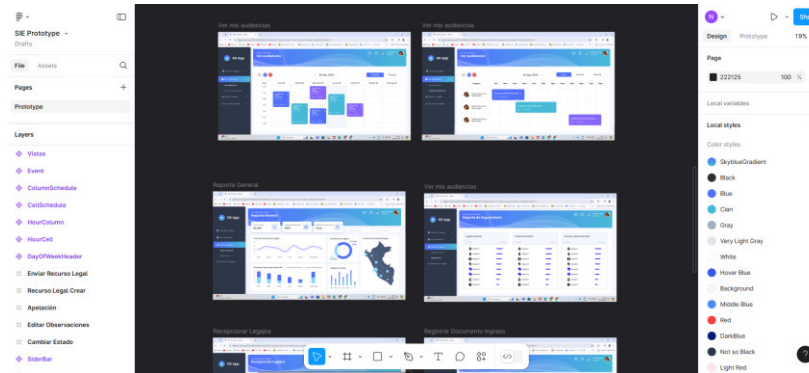
Capa de Procesamiento Inteligente (App Service 3: Backend Python). El backend en Python está orientado al procesamiento automatizado de documentos mediante OCR y LLM (modelo Llama 3.1.). Este servicio se divide en dos fases principales: extracción e identificación. El detalle del flujo de procesamiento corresponde al sprint 1 y se encuentra en la sección 5.3.3.1.

Otros servicios de Azure utilizados. La solución utiliza servicios de almacenamiento y gestión de datos proporcionados por Azure. La información estructurada se almacena en una base de datos SQL, mientras que los documentos digitales se gestionan mediante una cuenta de almacenamiento de archivos (Blob Storage). Asimismo, el sistema emplea un almacén de claves para la protección de credenciales y configuraciones sensibles, garantizando la seguridad de la información.

5.3.2.3. Prototipado. Para elaborar el diseño del sistema SIE se utilizó la plataforma de Figma, una herramienta muy versátil que permite de forma intuitiva la creación tanto de wireframes como prototipos de alta fidelidad como los que se muestran a continuación.

Figura 31

Prototipo de la solución.



Nota. Elaboración Propia.

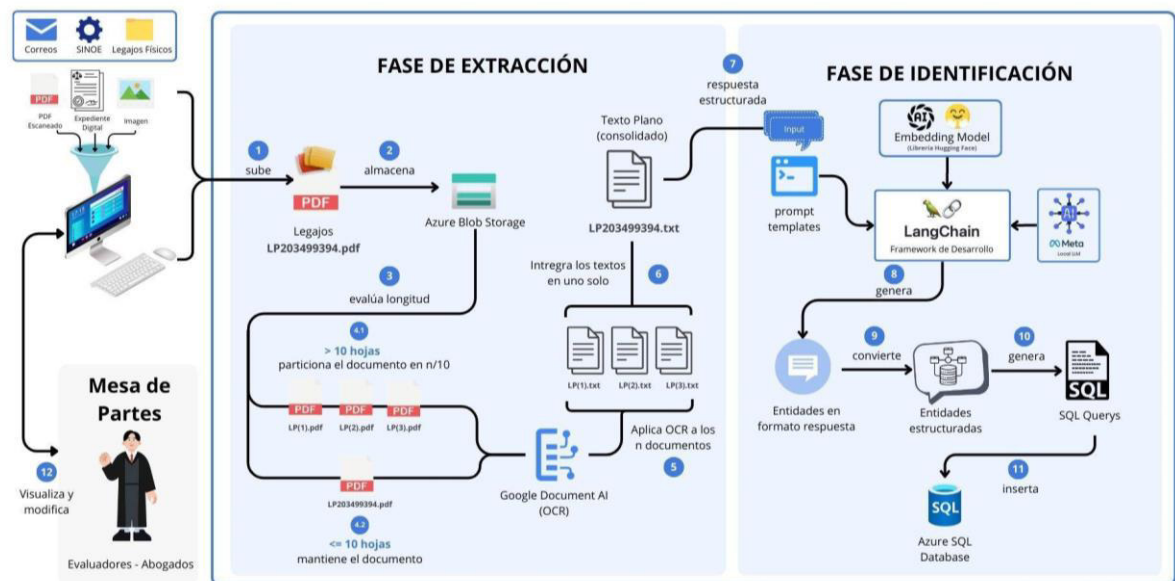
5.3.3. Implementación del nuevo sistema

5.3.3.1. Sprint 1: Servicio AI - Procesamiento de Documentos.

Flujo de procesamiento de documentos

Figura 32

Pipeline de Automatización con Enfoque Human-in-the-loop



Nota. Elaboración Propia.

El sistema es una API en Python para procesar documentos judiciales escaneados (PDF o imágenes provenientes de papel, sin texto digital), con enfoque human-in-the-loop, donde el usuario puede revisar y corregir la información antes de guardarla. La arquitectura se divide en dos fases:

- **Fase de extracción:**

- El usuario carga el documento y se almacena temporalmente en servicios de Microsoft (Azure Blob Storage).
- Si el documento tiene más de 10 páginas, se divide automáticamente para mejorar el procesamiento.
- Se aplica OCR con Document AI de Google para convertir la imagen en texto.
- Finalmente, los textos obtenidos se consolidan en un único archivo.

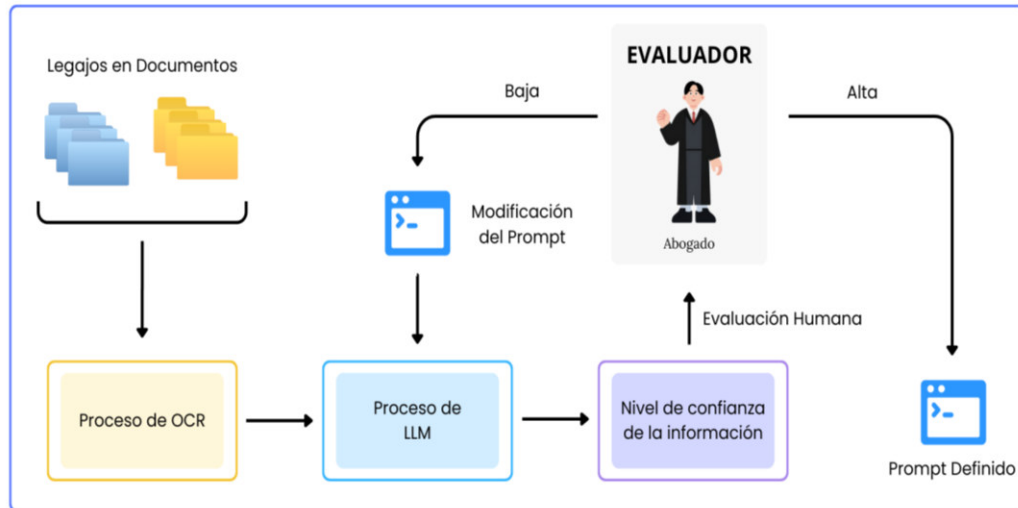
- **Fase de identificación con LLM**

- El texto se analiza usando prompts definidos y el framework LangChain.
- Se utiliza el modelo Llama 3.1 de Meta, ejecutado de forma local y configurado con los siguientes hiperparámetros: max_output_tokens: 8192, temperature: 0.2 y top_p: 0.95, lo que permite generar respuestas extensas, con baja aleatoriedad y alta coherencia semántica, garantizando mayor control y seguridad de la información procesada.
- El modelo extrae información jurídica organizada en:
 - Datos del documento.
 - Datos generales.
 - Partes procesales.
 - Audiencias y resultados.
- La salida se convierte en datos estructurados y consultas SQL.
- Se almacena en Azure SQL.
- Mesa de Partes, quien es un abogado, visualiza y corrige la información antes de confirmar el registro.

Etapas seguidas para el desarrollo

Figura 33

Etapas de Desarrollo



Nota. Elaboración Propia.

El flujo presentado en el diagrama se desarrolló siguiendo los principios de la metodología GenAIOps, la cual permite gestionar el ciclo de vida de aplicaciones basadas en modelos de inteligencia artificial generativa mediante procesos iterativos de desarrollo, evaluación y mejora continua.

A continuación, se describe cómo cada etapa de GenAIOps se refleja en el sistema propuesto:

- **Planificación y rediseño del sistema SIE:** En el marco de la reingeniería del sistema web SIE, se definió como objetivo mejorar la gestión de expedientes judiciales en la Procuraduría Pública del Sector Interior – MININTER mediante la automatización del procesamiento de documentos escaneados. Se identificó como problemática el alto tiempo y baja eficiencia del registro manual, por lo que se incorporaron tecnologías de OCR y modelos de lenguaje (LLM) para la extracción semántica de información jurídica. Se diseñó una arquitectura modular integrada al

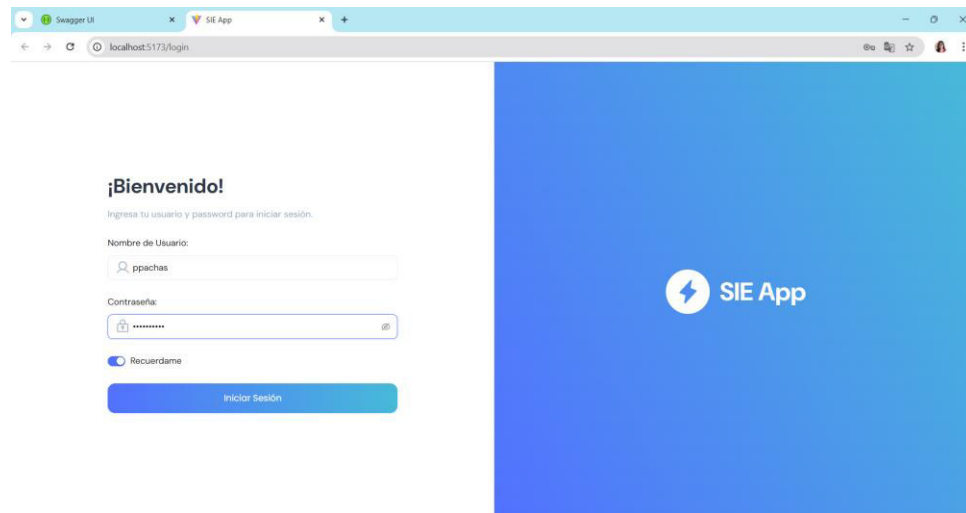
SIE y se adoptó el enfoque Human-in-the-Loop, garantizando supervisión experta en el proceso.

- **Pruebas, ajuste iterativo y validación:** Para validar la propuesta, se utilizaron 100 documentos judiciales del área penal con distribución no uniforme, simulando condiciones reales. Cada documento fue procesado automáticamente y evaluado por un abogado especialista mediante un instrumento de exactitud. Cuando los resultados no alcanzaban el nivel esperado, se ajustaban los prompts y se reprocesaban los casos, permitiendo una mejora progresiva del desempeño. Asimismo, ante cambios relevantes, se realizaron pruebas de validación y regresión para asegurar estabilidad y precisión.
- **Despliegue, monitoreo y mejora continua:** Al alcanzar un nivel adecuado de desempeño, el sistema reingenierizado fue preparado para su integración en el entorno operativo del SIE. Se estableció un proceso de monitoreo continuo, retroalimentación y evaluación permanente de precisión, confiabilidad y consistencia. De esta manera, la reingeniería no solo introduce automatización inteligente, sino un modelo evolutivo alineado con GenAIOps, orientado a mejorar de forma sostenible la gestión de expedientes judiciales.

5.3.3.2.Sprint 2: Módulo de Login y Configuraciones. El segundo sprint comprendió la ejecución de tres user stories que involucran las siguientes funcionalidades: inicio de sesión, autenticación a través de tokens e instalación de las dependencias necesarias tanto en Net 8 como en React. De esta manera se logró un inicio de sesión completamente funcional y una configuración completa de las cookies para el almacenamiento de información relevante del usuario como los tokens de autenticación.

Figura 34

Pantalla Inicio de Sesión.

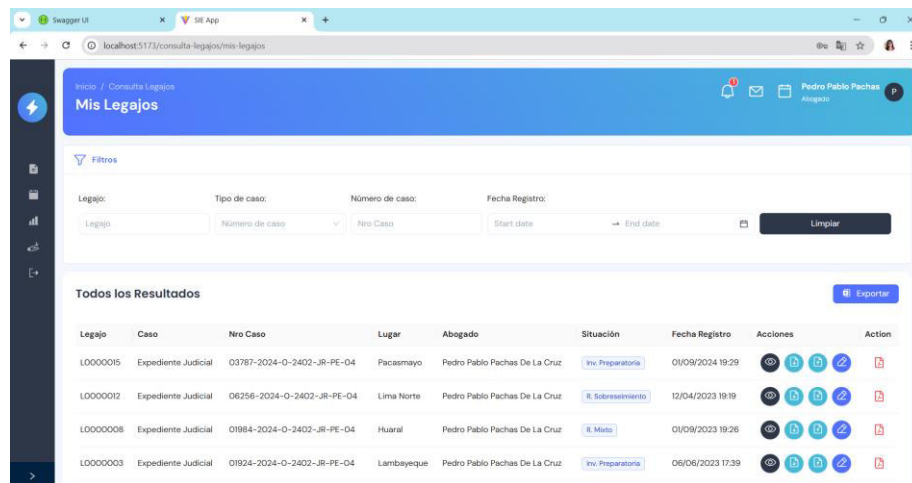


Nota. Elaboración Propia.

5.3.3.3.Sprint 3: Módulo de Consulta de legajos.

Figura 35

Pantalla Mis legajos.



Nota. Elaboración Propia.

El tercer sprint comprendió 19 user stories que abarcaban las funcionalidades básicas de las pantallas de Mis legajos y Todos los legajos. En estos apartados del sistema, el usuario puede buscar legajos a través de su código, el tipo de caso, el número de caso y

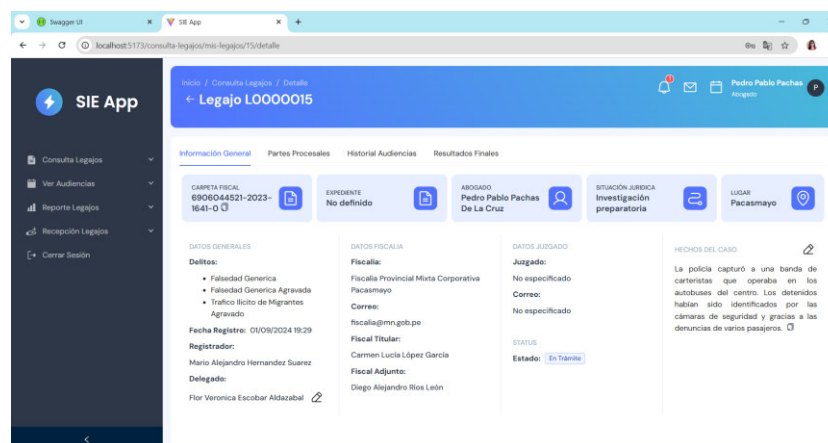
la fecha de registro. También puede exportar cada legajo en formato PDF o exportar el listado de todos los legajos de la página actual en formato Excel. Asimismo, por cada legajo el usuario cuenta con las siguientes opciones:

- Ver detalles
- Ver documentos de ingreso
- Ver documentos de salida
- Editar estado.

El botón de Ver detalle redirige a una página que posee 4 tabs. En cada tab se puede visualizar información relevante del caso judicial. Por ejemplo, en el tab de “Información General” se evidencia una subdivisión en otros 4 apartados, donde tenemos datos generales, datos de la fiscalía, datos del juzgado y hechos del caso.

Figura 36

Pantalla legajo - Información General.

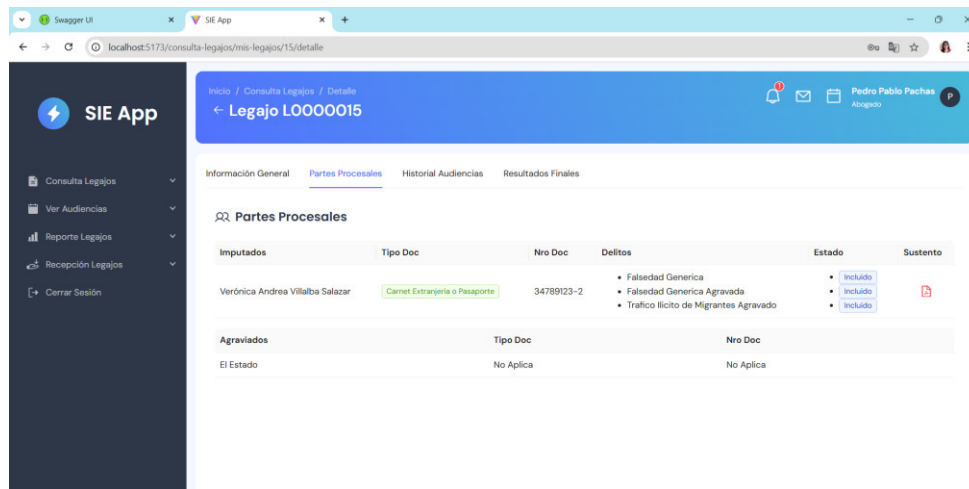


Nota. Elaboración Propia.

Respecto a la pestaña de “Partes Procesales”, vemos información de imputados y agraviados donde se especifica el tipo de documento y número de documento para cada uno de ellos. La principal diferencia entre la tabla de imputados y agraviados es que la tabla de imputados contiene más información relacionada con los delitos, el estado del delito y su sustento.

Figura 37

Pantalla legajo – Partes Procesales.

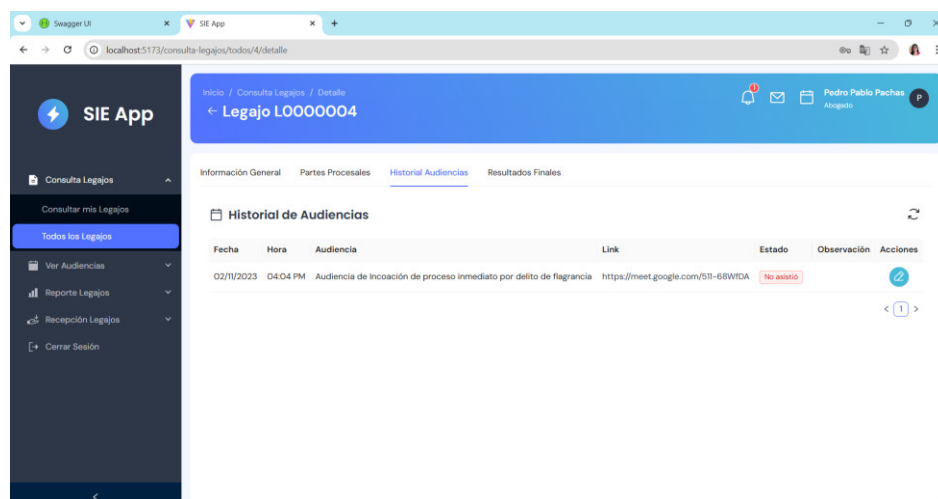


Nota. Elaboración Propia.

Siguiendo con la pestaña del “Historial de audiencias”, se resaltan los detalles de las audiencias en las que el abogado está trabajando. Esta sección proporciona detalles importantes sobre la audiencia, incluso si es en persona o virtual, y cualquier observación adicional que desee registrar.

Figura 38

Pantalla legajo – Historial de Audiencias.

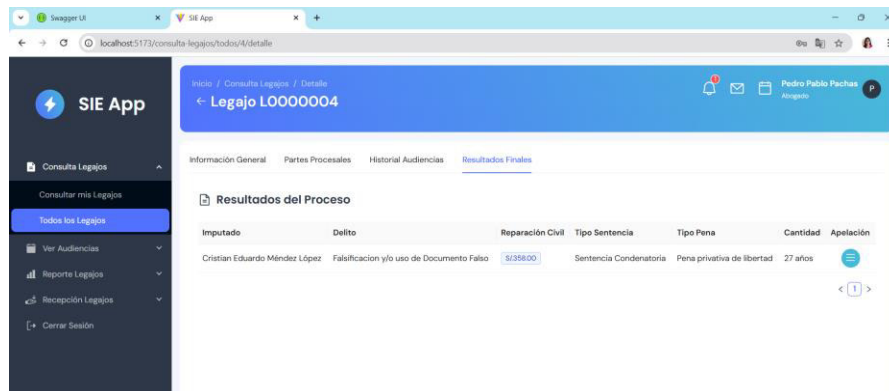


Nota. Elaboración Propia.

Finalmente, está la última pestaña de “Resultados del proceso”. Aquí se encuentra información de la sentencia emitida por cada delito de un imputado dentro del legajo.

Figura 39

Pantalla legajo – Resultados Finales.

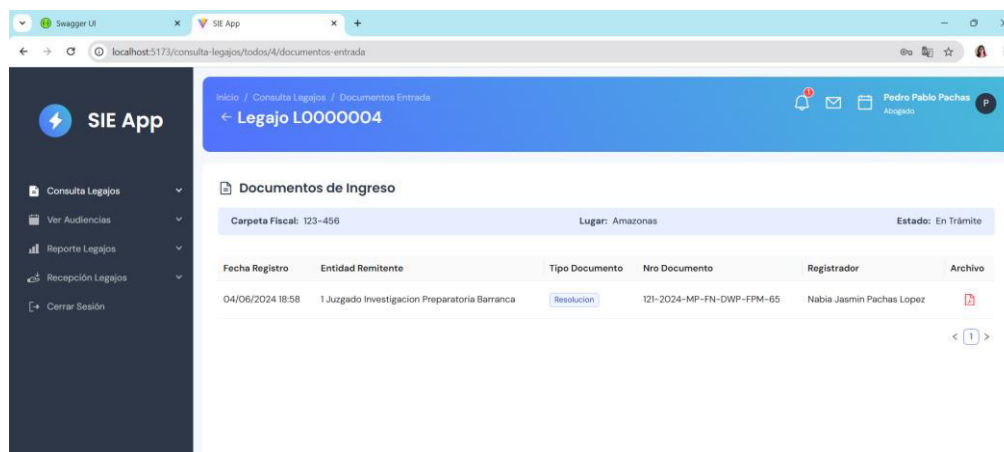


Nota. Elaboración Propia.

Por otro lado, la opción de “Ver documentos de ingreso” permite visualizar todos los documentos recepcionados por la procuraduría relacionados a ese caso penal en específico. Por cada documento de ingreso, se visualizan datos como la fecha del registro, la entidad remitente, el tipo de documento que está siendo recepcionado, el número de documento, el registrador y el archivo para visualizarlo en caso se considere necesario.

Figura 40

Pantalla legajo – Documentos de Ingreso



Nota. Elaboración Propia.

Asimismo, la opción “Ver documentos de salida” permite acceder a la pantalla “Documentos de salida” (véase Figura 41). Un documento de salida, o recurso legal, representa una respuesta a uno o varios documentos de entrada. Actualmente, el sistema permite elaborar automáticamente diversos tipos de documentos de salida, como la Constitución en actor civil, el formato de archivo, entre otros.

Una vez seleccionada la opción “Crear”, el documento se registra como un nuevo elemento en la tabla de documentos de salida, habilitándose las opciones para descargarlo o enviarlo por correo electrónico a distintos destinatarios. En cuanto a la funcionalidad de envío, el usuario puede seleccionar destinatarios predeterminados, como fiscalías o juzgados, o bien ingresar de manera personalizada una dirección de correo electrónico.

Figura 41

Pantalla legajo – Documentos de Salida

Fecha Creación	Abogado	Tipo Documento	Destinatarios	Archivo	Enviar
14/10/2024 20:10	Nabia Jasmin Pachas Lopez	Actor Civil			
14/10/2024 21:17	Nabia Jasmin Pachas Lopez	Actor Civil			
14/10/2024 21:35	Nabia Jasmin Pachas Lopez	Actor Civil	- Juzgado Penal Unipersonal Trujillo - Otro de prueba - Juzgado Penal Unipersonal Trujillo - Otro de prueba		
22/10/2024 18:33	Pedro Pablo Pachas De La Cruz	Actor Civil	Nabia Jasmin Pachas Lopez		

Nota. Elaboración Propia.

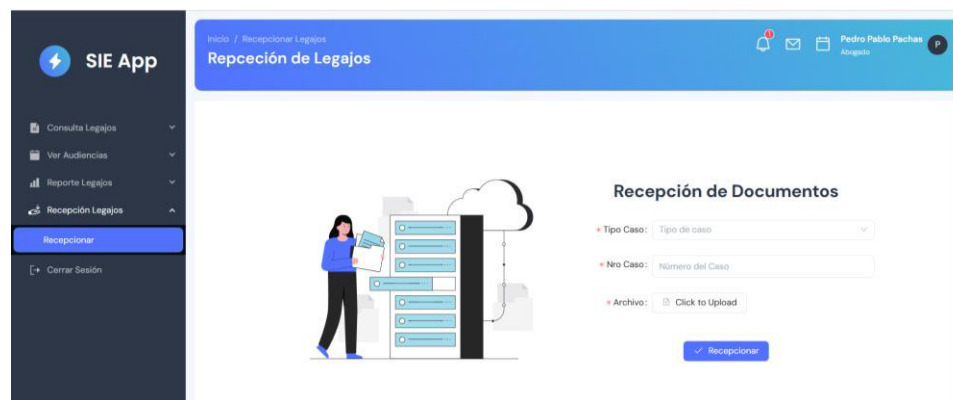
5.3.3.4.Sprint 4: Módulo de Recepción de legajos. El cuarto sprint corresponde al módulo de Recepción de legajos. Como se ha mencionado anteriormente, el sistema gestiona tanto documentos de entrada como de salida; en el caso de los documentos de entrada, estos son recepcionados a través del módulo de Recepción de Legajos (véase Figura 42). En esta pantalla, se solicita al usuario ingresar el tipo de caso, su número y el archivo del documento en formato PDF.

Una vez cargado el archivo y seleccionada la opción “Recepcionar”, el sistema verifica internamente en la base de datos si el documento ya ha sido registrado, con el fin de determinar si corresponde a la creación de un nuevo legajo o a la incorporación de un documento adicional en uno existente. Posteriormente, se realiza una llamada a la API desarrollada en Python para efectuar la extracción de las entidades relevantes.

Finalmente, una vez concluido este proceso y según el resultado de la verificación inicial, el sistema redirige al usuario a la pantalla correspondiente.

Figura 42

Pantalla Recepción legajos



Nota. Elaboración Propia.

Al confirmarse que se trata de un nuevo legajo, el sistema redirige al usuario a una nueva pantalla (véase Figura 43), la cual se encuentra organizada en cuatro secciones desplegables (collapsers), donde se presenta la información contenida en el documento. Como se observa en dicha pantalla, las entidades relevantes ya han sido extraídas y

completadas automáticamente, lo que permite agilizar el procesamiento de la información. No obstante, estos datos pueden ser modificados por el usuario en caso sea necesario.

Figura 43

Pantalla Nuevo legajo

The screenshot displays the 'Nuevo legajo' (New Case) screen within the SIE App. On the left, a dark sidebar contains navigation icons and labels: 'Consulta Legajos', 'Ver Audiencias', 'Reporte Legajos', 'Recepción Legajos', and 'Cerrar Sesión'. The main content area has a blue header with 'Inicio / Recepción Legajos / 1' and a user profile for 'Kevin Miguel Ortiz Admin'. Below the header, the title 'Legajo' is followed by a 'Número Legajo' field containing 'L0000004' and a 'Guardar Datos' button. A message indicates that data was automatically identified from the document. The form is organized into expandable sections: 'Datos del documento' (showing 'Clase de Documento' as 'Disposicion Fiscal', 'Nro de Documento' as '267-2024-MA-FN-DPP-FPM-DZ', 'Tipo Remitente' as 'Fiscalia', and 'Remitante' as 'Fiscalia Provincial Mixta Corporativa La E...'), 'Datos Generales del Caso', and 'Partes Procesales'.

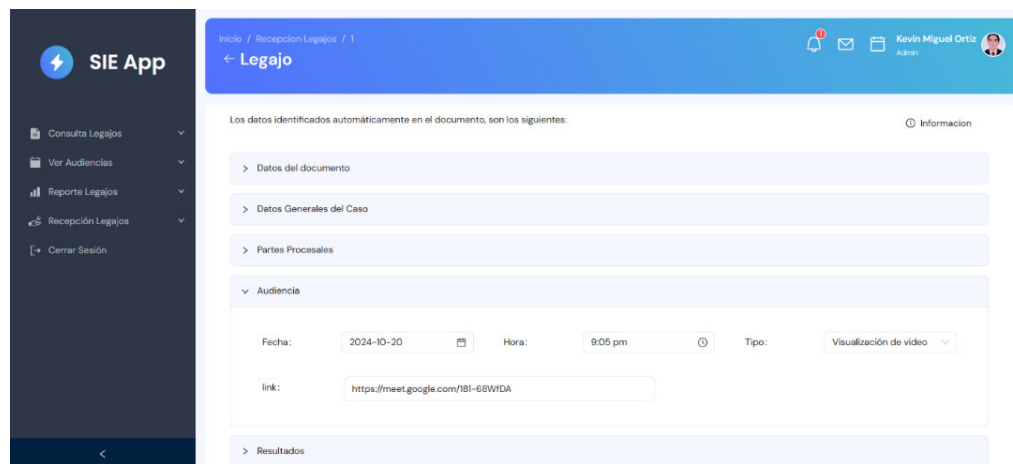
Nota. Elaboración Propia.

En caso se trate de un legajo existente, el sistema redirige a la pantalla “Adicionar documento” (véase Figura 44) la cual presenta una estructura similar, organizada también en cuatro secciones desplegables que contienen la misma información. La principal diferencia entre las opciones de nuevo legajo y adicionar documento radica en la gestión de los datos generales del caso y de las partes procesales. En la opción de nuevo legajo, se muestra información relacionada con la fiscalía ya que, en las etapas iniciales del proceso penal, se define la fiscalía a cargo y el número de carpeta fiscal.

Por otro lado, en esta misma opción se registra la información de los agraviados, la cual se almacena una única vez, debido a que los documentos posteriores hacen referencia a estos datos. En consecuencia, dicha información sólo se encuentra disponible en la pantalla de nuevo legajo. Fuera de estas diferencias, el resto de los campos mantiene la misma estructura y contenido en ambos apartados.

Figura 44

Pantalla Adicionar legajo



Nota. Elaboración Propia.

5.3.3.5.Sprint 5: Módulo de Reporte de legajos. El quinto sprint hace referencia al reporte de legajos, este apartado es útil para el procurador, dado que este módulo tiene como objetivo visualizar los estados en los que se encuentran los legajos, también se encuentra la opción de hacer uso de filtros por año, mes, delito. En lo que respecta gráficos, permite apreciar su evolución temporal, en qué etapa se encuentran mayor cantidad de legajos, cuanto representa dicha cantidad, su localización geográfica, entre otros aspectos.

Figura 45

Pantalla Reporte General

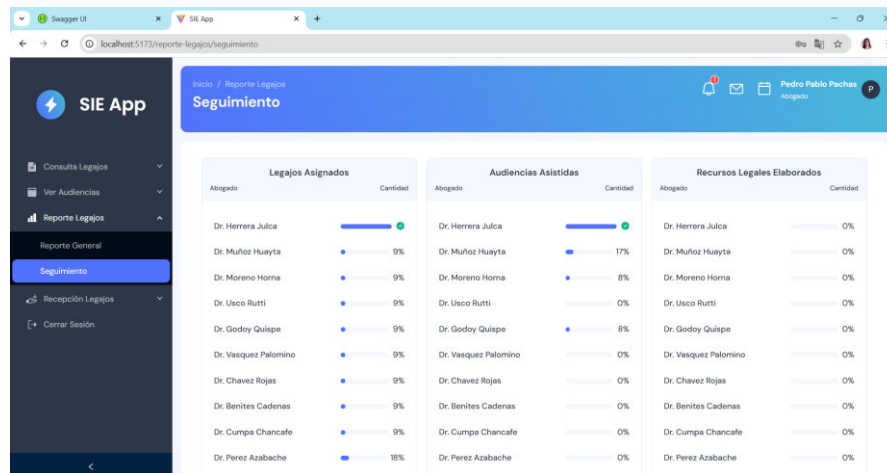


Nota. Elaboración Propia.

En la pantalla anterior se puede observar de forma conjunta y global el estado de los legajos. Si el procurador desea realizar un seguimiento más exhaustivo por parte del abogado, se ha creado la vista de Seguimiento (véase Figura 46), en la que se destacan métricas y estadísticas de gran relevancia. Una de las políticas institucionales establece que un abogado debe presentar y completar una serie de legajos y documentos mensualmente. Por ende, esta pantalla proporciona una información más exacta en relación con los legajos asignados, las audiencias asistidas y los recursos legales generados.

Figura 46

Pantalla Seguimiento



Nota. Elaboración Propia.

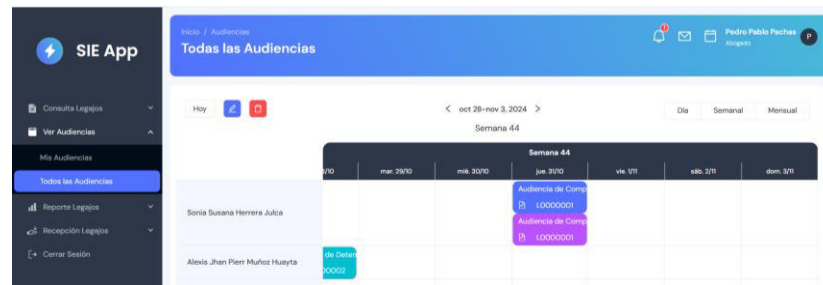
5.3.3.6.Sprint 6: Módulo de Audiencias. El módulo de audiencias ofrece una perspectiva más extensa de la planificación y programación mediante un diagrama de Gantt o un calendario. Es crucial tener en cuenta que se ofrecen dos vistas que varían dependiendo del tipo de usuario que accede. Por ejemplo, si el usuario corresponde a un procurador, verá la pantalla de Audiencias, que se refiere a las audiencias de cada abogado en una fecha determinada.

Esto ofrece al procurador una perspectiva más extensa para supervisar la asistencia de los abogados y participar en alguna si se requiere. Como se puede apreciar en la Figura 47, es posible realizar una perspectiva diaria, semanal y mensual del diagrama de Gantt. Además, se tiene la opción de modificar las mismas, esta versión permite cambiar el

intervalo de tiempo de la audiencia y la fecha que se estaba evaluando, incluso se puede suprimir y crear una audiencia cuando sea necesario. Así logra un mayor control sobre las audiencias que se están realizando.

Figura 47

Pantalla Audiencias

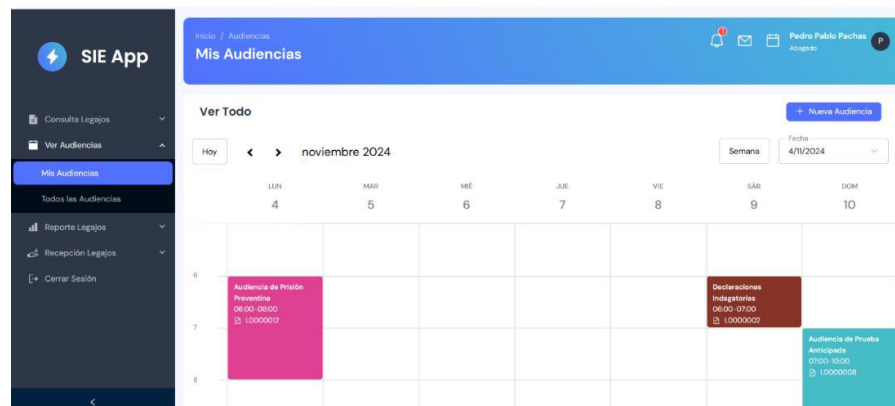


Nota. Elaboración Propia.

Por otro lado, se tiene la pantalla de mis audiencias. El contenido que posee es beneficioso para la visualización del abogado, ya que le permite realizar un seguimiento de las audiencias programadas accediendo a la plataforma. Al igual que en la pantalla anterior, estas se pueden observar diaria, semanal y mensualmente. La edición se realiza eligiendo cada uno de los recuadros y ajustando el rango de fechas arrastrándolo y extendiéndolo según sea necesario. Además, se cuenta con un botón que permite agregar nuevas audiencias al calendario manualmente.

Figura 48

Pantalla Mis Audiencias



Nota. Elaboración Propia.

Capítulo VI

Análisis de Resultados

6.1. Análisis de Resultados para la población 1

6.1.1. Análisis de Confiabilidad

El presente análisis de confiabilidad se realiza con la finalidad de evaluar la consistencia interna del instrumento aplicado en la Población 1, correspondiente a la población de abogados de la Procuraduría Pública.

La confiabilidad del instrumento se evalúa mediante el coeficiente Alfa de Cronbach, el cual permite determinar si los ítems que conforman cada dimensión presentan coherencia estadística.

6.1.1.1. Dimensión 1: Recepción y Consulta de Expedientes Judiciales. La Tabla 16 presenta los valores del coeficiente Alfa de Cronbach obtenidos para la dimensión “Recepción y Consulta de Expedientes Judiciales”, tanto en la etapa de pretest como del posttest. En el pretest, el valor del Alfa de Cronbach (0.909) y del Alfa estandarizado (0.914) indican una alta consistencia interna entre los cuatro ítems que conforman esta dimensión.

Tabla 16

Alfa de Cronbach para la Dimensión 1

	Alfa de Cronbach	Alfa de Cronbach Estandarizado	Nº Elementos
Pretest	0.909	0.914	4
Posttest	0.956	0.958	4

Nota. Elaboración Propia.

En el posttest, los valores de confiabilidad aumentan significativamente, alcanzando un Alfa de Cronbach de 0.956 y un Alfa estandarizado de 0.958. Este incremento sugiere una mayor homogeneidad en las respuestas de los participantes tras la implementación del nuevo sistema.

6.1.1.2.Dimensión 2: Elaboración y Envío de Recursos Legales. En la Tabla 17 se muestran los resultados del análisis de confiabilidad para la dimensión “Elaboración y Envío de Recursos Legales”. En el pretest, el Alfa de Cronbach obtenido (0.883) evidencia una buena consistencia interna entre los ítems.

Tabla 17

Alfa de Cronbach para la Dimensión 2

	Alfa de Cronbach	Alfa de Cronbach Estandarizado	Nº Elementos
Pretest	0.883	0.883	4
Posttest	0.910	0.916	4

Nota. Elaboración Propia.

Posteriormente, en el posttest, el Alfa de Cronbach se incrementa a 0.910 y el Alfa estandarizado a 0.916, lo que refleja una mejora en la coherencia de las respuestas.

6.1.1.3.Dimensión 3: Planificación y Participación en Audiencias. La Tabla 18 presenta los valores del Alfa de Cronbach correspondientes a la dimensión “Planificación y Participación en Audiencias”. En el pretest, se obtuvo un coeficiente de 0.947, lo cual indica un nivel de confiabilidad muy alto entre los cuatro ítems evaluados.

Tabla 18

Alfa de Cronbach para la Dimensión 3

	Alfa de Cronbach	Alfa de Cronbach Estandarizado	Nº Elementos
Pretest	0.947	0.947	4
Posttest	0.935	0.936	4

Nota. Elaboración Propia.

En el posttest, el Alfa de Cronbach se mantiene en niveles elevados (0.935), evidenciando que la consistencia interna del instrumento se conserva tras la implementación del nuevo sistema.

6.1.1.4. Dimensión 4: Revisión y Seguimiento de Expedientes. En la Tabla 19 se muestran los resultados del análisis de confiabilidad para la dimensión “Revisión y Seguimiento de Expedientes”. En el pretest, el Alfa de Cronbach alcanzó un valor de 0.900, lo que indica una alta consistencia interna entre los ítems de esta dimensión.

Tabla 19

Alfa de Cronbach para la Dimensión 4

	Alfa de Cronbach	Alfa de Cronbach Estandarizado	Nº Elementos
Pretest	0.900	0.905	4
Posttest	0.945	0.948	4

Nota. Elaboración Propia.

En el posttest, los valores de confiabilidad se incrementan hasta un Alfa de Cronbach de 0.945 y un Alfa estandarizado de 0.948.

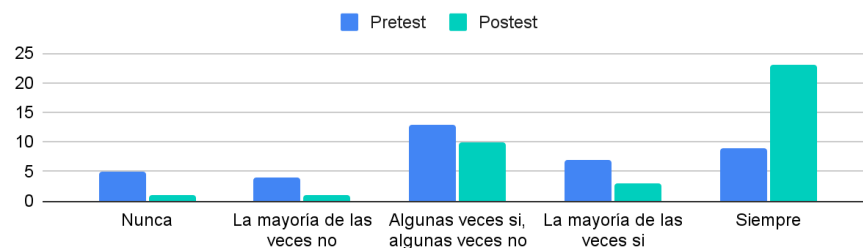
6.1.2. Análisis Estadístico Descriptivo

6.1.2.1. Análisis por ítems. El cuestionario utilizado para la recolección de datos de la población 1 consta de 16 preguntas divididas en 4 secciones. El detalle de los resultados obtenidos se muestra a continuación:

Ítem 1: Los expedientes judiciales que me asignan diariamente contienen la información completa en el sistema

Figura 49

Gráfico de frecuencia de la Pregunta n°1 por pretest y posttest



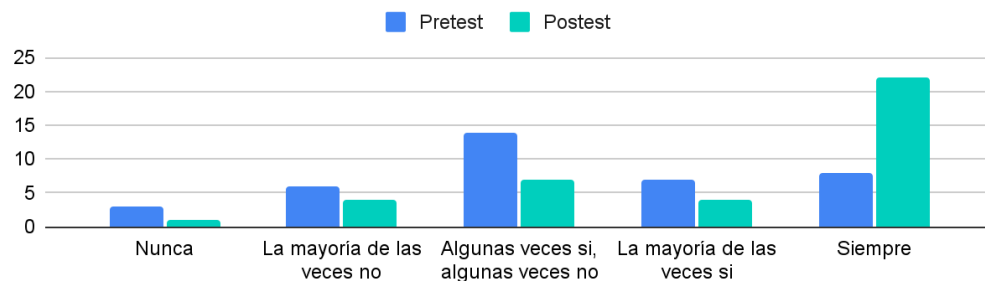
Nota. Elaboración Propia.

Como se puede apreciar en la Figura 49, en el pretest, la mayoría de los encuestados se centran en dos opiniones: 9 personas piensan que siempre hay disponibilidad de información, mientras que 13, sienten que solo algunas veces. Asimismo, 4 personas opinan que la mayoría de las veces no contiene la información necesaria. Por otro lado, para el posttest, la mayoría considera que siempre se encuentra la información con una frecuencia de 23, y una minoría opina que nunca y la mayoría de las veces no, ambos con una cantidad 1.

Ítem 2: En el sistema SIE no he encontrado errores en la información cargada sobre los expedientes judiciales

Figura 50

Gráfico de frecuencia de la Pregunta n°2 por pretest y posttest



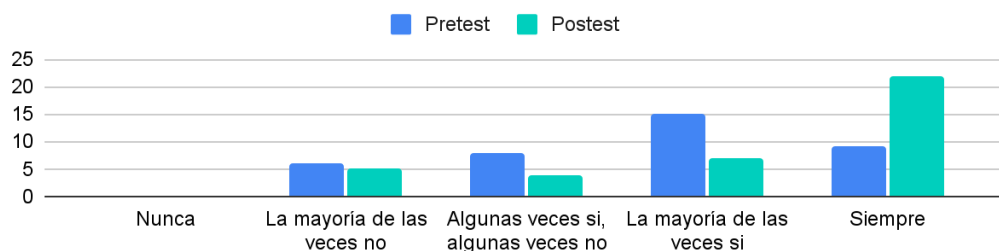
Nota. Elaboración Propia.

Como se puede apreciar en la Figura 50, en el pretest, la mayoría de los encuestados se centran en dos opiniones: 8 personas consideran que siempre tienen la percepción que no se encuentran errores en el sistema, mientras que 14, sienten que algunas veces encuentran errores y en otras no. Asimismo, 3 personas opinan que nunca han sentido que no hay errores en el sistema. Por otro lado, para el posttest, la mayoría considera que siempre ha sentido que el sistema no presenta errores con una frecuencia de 22.

Ítem 3: No se me ha dificultado acceder a la información subida sobre los expedientes judiciales en el sistema SIE

Figura 51

Gráfico de frecuencia de la Pregunta n°3 por pretest y posttest



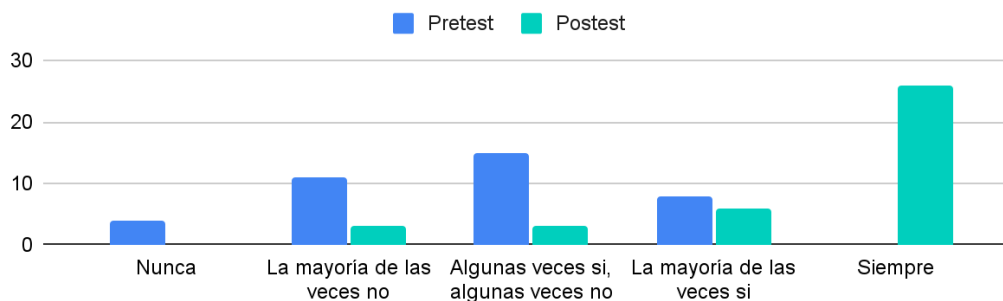
Nota. Elaboración Propia.

Como se puede apreciar en la Figura 51, en el pretest, la mayoría de los encuestados se centran en dos opiniones: 9 personas consideran nunca haber tenido problemas en acceder a la información, mientras que 15, sienten que la mayoría de las veces no han tenido problemas. Asimismo, 8 personas opinan que algunas veces ha tenido dificultad para acceder a la información. Por otro lado, para el posttest, la mayoría no ha presentado dificultades en el acceso de información con una frecuencia de 22. Tanto en el posttest como el pretest, ninguna contó con la opinión de nunca.

Ítem 4: El proceso de digitación y carga de la información de un expediente judicial en el sistema se realiza de manera rápida y toma muy poco tiempo

Figura 52

Gráfico de frecuencia de la Pregunta n°4 por pretest y posttest



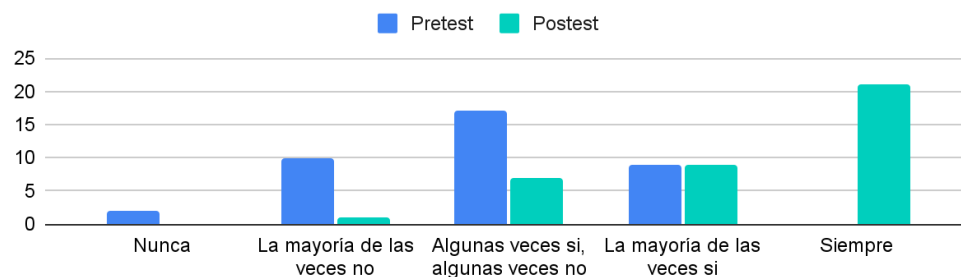
Nota. Elaboración Propia.

Como se puede apreciar en la Figura 52, en el pretest, la mayoría de los encuestados se centran en dos opiniones, 11 personas piensan que la mayoría de las veces el sistema no toma mucho tiempo en cargar la información, mientras que 15, sienten que algunas veces encuentran demora de tiempo y en otras no. Una minoría opina que nunca han sentido que el sistema carga la información a tiempo, representada por 4 personas. Por otro lado, para el posttest, la mayoría considera que el sistema no presenta errores con una frecuencia de 26. Una apreciación importante para resaltar es que ninguna persona eligió la opción de nunca en el posttest.

Ítem 5: El tiempo promedio que me toma elaborar un recurso legal (formato de archivo PSI, apersonamiento, etc.) no demora más de 10 minutos

Figura 53

Gráfico de frecuencia de la Pregunta n°5 por pretest y posttest



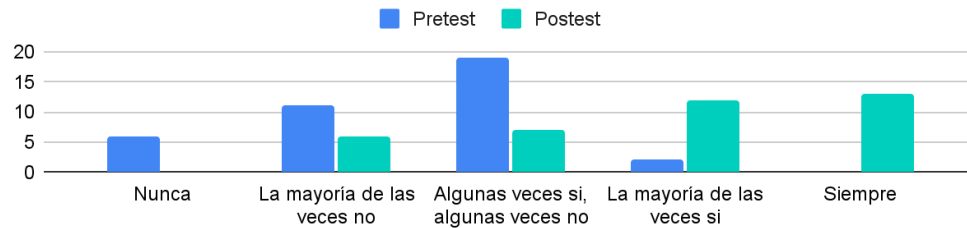
Nota. Elaboración Propia.

Como se puede apreciar en la Figura 53, en el pretest, la mayoría de encuestados se centran en dos opiniones: 10 personas piensan que toma más de 10 minutos realizar un recurso legal, mientras que 17, sienten que algunas veces emplean más de 10 minutos y en otras no. Por otro lado, para el posttest, la mayoría considera que no demora más de 10 minutos en realizar un recurso legal con una frecuencia de 21, y una minoría opina que la mayoría de las veces no, representado por 1 persona. Adicionalmente ninguna persona seleccionó la opción de nunca.

Ítem 6: La carga de recursos legales, que tengo para elaborar durante el día, no me parece excesivo o pesado

Figura 54

Gráfico de frecuencia de la Pregunta n°6 por pretest y posttest



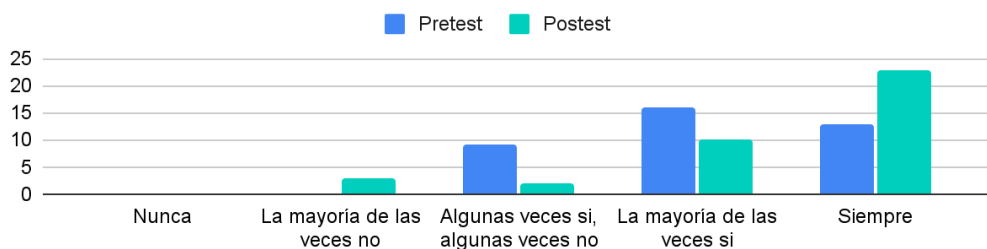
Nota. Elaboración Propia.

Como se puede apreciar en la Figura 54, en el pretest, la mayoría de encuestados se centran en dos opiniones: 11 personas consideran tener una carga excesiva, mientras que 19, sienten que algunas veces les parece que tienen una carga excesiva y en otras no. Por otro lado, en lo que respecta al posttest, la mayoría se inclina en dos opciones: 12 personas consideran no tener una carga excesiva la mayoría de las veces, mientras que 13, opinan que siempre.

Ítem 7: La tasa de aceptación de los recursos legales que envió a otras entidades es buena

Figura 55

Gráfico de frecuencia de la Pregunta n°7 por pretest y posttest



Nota. Elaboración Propia.

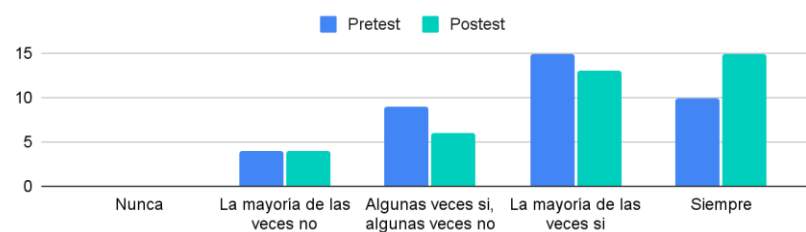
Como se puede apreciar en la Figura 55, en el pretest, la mayoría de encuestados se centran en dos opiniones: 16 personas piensan que la mayoría de las veces tienen la tasa de aceptación es buena, mientras que 13, sienten que siempre es buena. Una minoría opina

que a veces la tasa sí es buena y otras veces no lo es, representada por 9 personas. Por otro lado, para el postest, la mayoría siempre considera que la tasa de aceptación es buena con una frecuencia de 23.

Ítem 8: No he encontrado errores en la información y la redacción de los recursos legales posteriormente a su envío a otra entidad

Figura 56

Gráfico de frecuencia de la Pregunta n°8 por pretest y postest



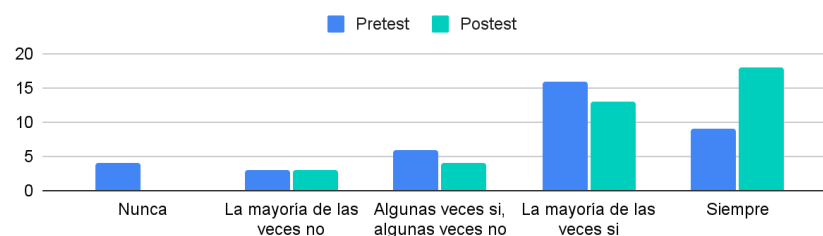
Nota. Elaboración Propia.

Como se puede apreciar en la Figura 56, en el pretest, la mayoría de encuestados se centran en dos opiniones: 10 personas consideran que siempre se han identificado errores en información o redacción, mientras que 15, sienten que la mayoría de las veces no encuentran errores. Por otro lado, para el postest, la mayoría se inclina por dos opciones: 15 personas no han sentido presencia de errores de información o redacción, y 13 personas dicen que la mayoría de las veces no encuentran errores.

Ítem 9: Actualmente, las audiencias programadas me son comunicadas de manera oportuna

Figura 57

Gráfico de frecuencia de la Pregunta n°9 por pretest y postest



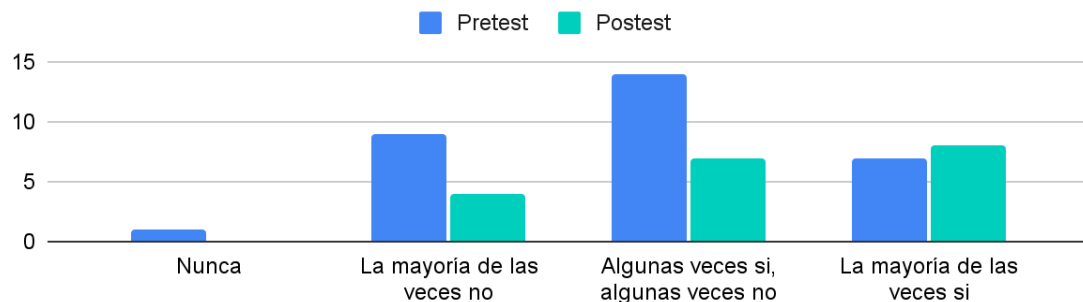
Nota. Elaboración Propia.

Como se puede apreciar en la Figura 57, en el pretest, la mayoría de los encuestados se centran en dos opiniones: 9 personas consideran que siempre existe una comunicación oportuna, mientras que 16, sienten que la mayoría de las veces es oportuna. Una minoría opina que la mayoría de las veces la comunicación no es oportuna, representada por 3 personas. Por otro lado, para el postest, la mayoría se inclina por dos opciones: 18 personas siempre sienten que la comunicación es oportuna, y 13 personas piensan que la mayoría de las veces si es oportuna.

Ítem 10: No he tenido problemas para acceder a la información de las audiencias (fecha, hora, modalidad, lugar o enlace de videoconferencia)

Figura 58

Gráfico de frecuencia de la Pregunta n°10 por pretest y postest



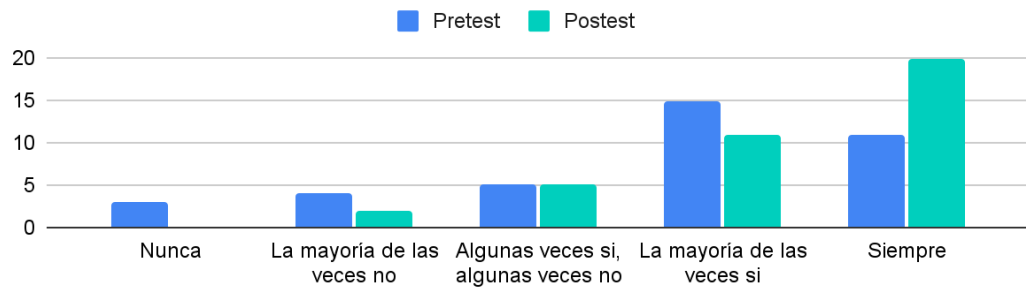
Nota. Elaboración Propia.

Como se puede apreciar en la Figura 58, en el pretest, se interpreta que la mayoría de los encuestados se centran en dos opiniones: 14 personas piensan que algunas veces no tienen problemas y otras veces sí, mientras que 9, sienten que la mayoría de las veces si han tenido problemas. Por otro lado, para el postest, la mayoría se inclina por dos opciones: 19 personas consideran que el sistema no presenta problemas en el acceso de audiencias, mientras que 8, opinan que la mayoría de las veces el sistema no presenta problemas.

Ítem 11: He asistido a las audiencias que he tenido asignadas sin ningún tipo de inconveniente

Figura 59

Gráfico de frecuencia de la Pregunta n°11 por pretest y posttest



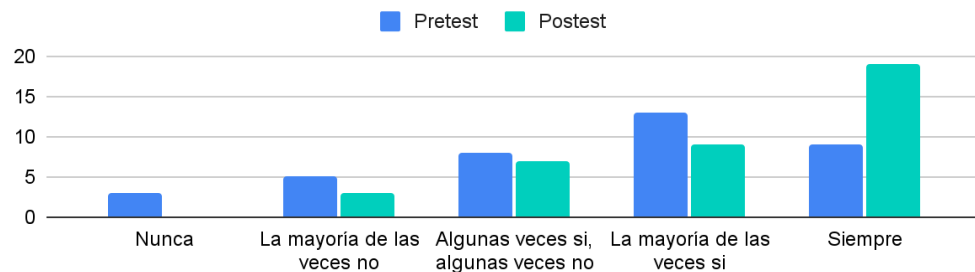
Nota. Elaboración Propia.

Como se puede apreciar en la Figura 59, en el pretest, la mayoría de los encuestados se centran en dos opiniones: 11 personas piensan que siempre han podido asistir a las audiencias sin problemas, mientras que 15, sienten que la mayoría de las veces han podido acceder sin inconvenientes. Por otro lado, para el posttest, la mayoría se inclina por dos opciones: 20 personas consideran que siempre han podido acceder a las audiencias sin ningún problema, mientras que 11, opinan que la mayoría de las veces.

Ítem 12: Durante mi participación en las audiencias no se me ha dificultado encontrar en el sistema la información del caso judicial para sustentar mi postura jurídica

Figura 60

Gráfico de frecuencia de la Pregunta n°12 por pretest y posttest



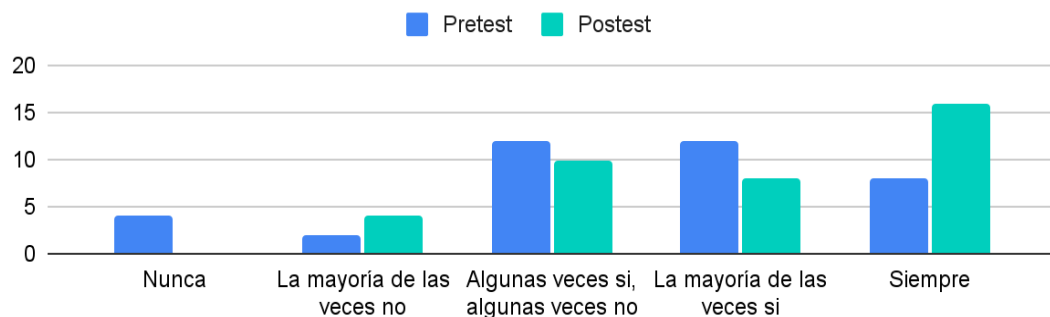
Nota. Elaboración Propia.

Como se puede apreciar en la Figura 60, en el pretest, se interpreta que la mayoría de los encuestados se centran en dos opiniones: 9 personas consideran no se ha dificultado identificar la información, mientras que 13, sienten que la mayoría de las veces no han encontrado dificultades. Por otro lado, para el postest, 19 personas consideran que no han presentado problemas en acceder a la información necesaria para su sustento, mientras que 9, opinan que la mayoría de las veces. Una minoría opina que la mayoría de las veces no, representado por 3 personas.

Ítem 13: He podido realizar la identificación de tendencias o patrones en los expedientes sin problemas

Figura 61

Gráfico de frecuencia de la Pregunta n°13 por pretest y postest



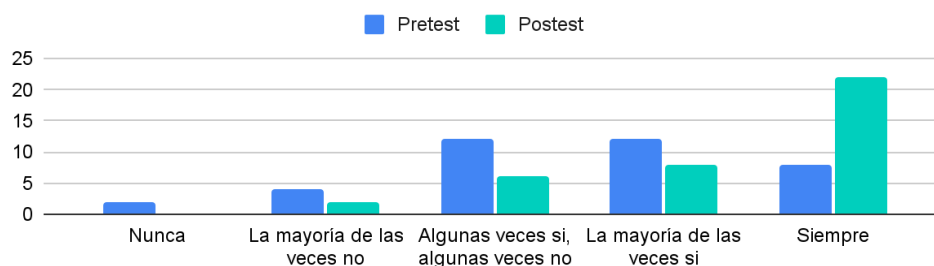
Nota. Elaboración Propia.

Como se puede apreciar en la Figura 61, en el pretest, la mayoría de encuestados se centran en dos opiniones: algunas veces no han podido identificar dichas tendencias y la mayoría de las veces sí, ambas representan una cantidad de 12. Por otro lado, para el postest, 16 personas afirman que siempre han podido realizar la identificación de tendencias en sus expedientes, mientras que 10, mencionan algunas veces sí han podido y algunas veces no.

Ítem 14: He podido realizar una revisión del estado de los expedientes de manera rápida y oportuna a través del sistema

Figura 62

Gráfico de frecuencia de la Pregunta n°14 por pretest y posttest



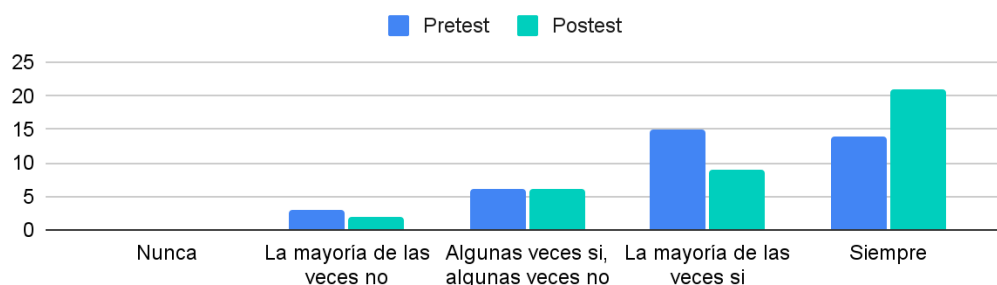
Nota. Elaboración Propia.

Como se puede apreciar en la Figura 62, en el pretest, se interpreta que la mayoría de las personas se centran en dos opiniones: algunas veces no han podido realizar una revisión del estado de manera rápida y la mayoría de las veces sí, ambas representan una cantidad de 12. Una minoría opina que nunca han realizado una revisión rápida, representada por 2 personas. Por otro lado, para el posttest, la mayoría se inclina por dos opciones, donde siempre se ha podido realizar una revisión rápida y oportuna de los estados con una frecuencia de 16, y algunas veces sí, representado por 10 personas.

Ítem 15: He podido informar a mi supervisor mi avance con respecto a la gestión de los expedientes judiciales que me han asignado

Figura 63

Gráfico de frecuencia de la Pregunta n°15 por pretest y posttest



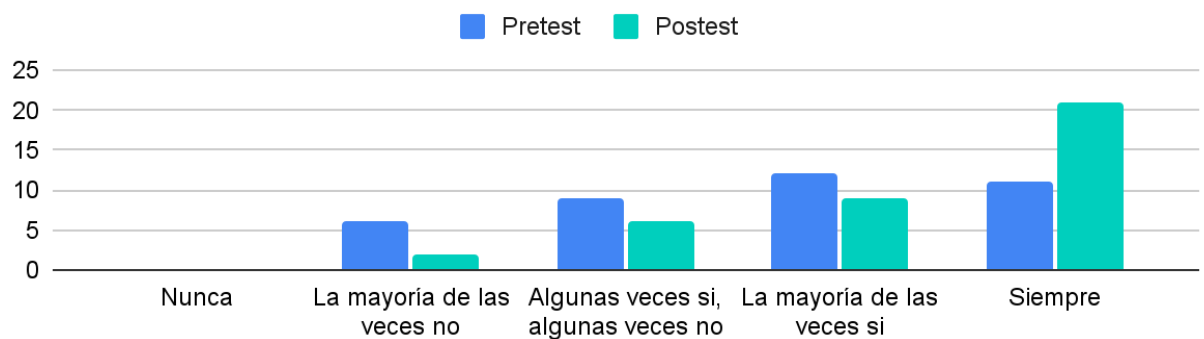
Nota. Elaboración Propia.

Como se puede apreciar en la Figura 63, en el pretest, la mayoría de los encuestados se centran en dos opiniones: 14 personas piensan que siempre pueden informar a su supervisor el avance obtenido, mientras que 15, sienten que algunas veces pueden informar y en otras en las que no. Una minoría opina que la mayoría de las veces no ha podido informar su avance, representada por 3 personas. Por otro lado, en lo que respecta el posttest, la mayoría se inclina a la opción de que siempre pueden informar con su supervisor su avance con una frecuencia de 21, y una minoría opina que la mayoría de las veces no, representado por 2 personas.

Ítem 16: En el sistema, he podido acceder a los puntos clave de cada expediente judicial para hacer el seguimiento respectivo

Figura 64

Gráfico de frecuencia de la Pregunta n°16 por pretest y posttest



Nota. Elaboración Propia.

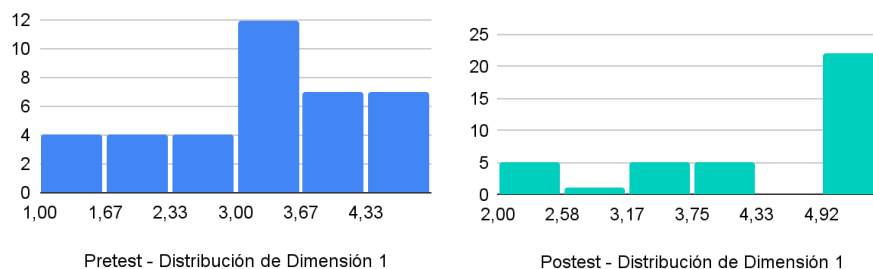
Como se puede apreciar en la Figura 64, en el pretest, la mayoría de los encuestados se centran en dos opiniones: 11 personas piensan que siempre tienen acceso a información clave para el seguimiento respectivo, mientras que 12, sienten que la mayoría de las veces cuentan con acceso. Por otro lado, para el posttest, la mayoría se inclina a la opción de que siempre pueden visualizar puntos claves en los expedientes para el seguimiento con una frecuencia de 21, y una minoría opina que la mayoría de las veces no, representado por 2 personas.

6.1.2.2. Análisis por dimensiones

Dimensión 1 (D1): Recepción y Consulta de Expedientes Judiciales

Figura 65

Histograma de la Dimensión 1



Nota. Elaboración Propia.

Tabla 20

Medidas clave de la Dimensión 1

Dimensión 1	Pretest	Posttest
Media	3,25	4,243421
Mediana	3,25	5
Desv. estándar	0,9966159	1,0241787
Varianza	0,993	1,049
Asimetría	-0,44	-0,996
Curtosis	-0,607	-0,423
Rango	3,5	3
Mínimo	1,25	2
Máximo	4,75	5

Nota. Elaboración Propia.

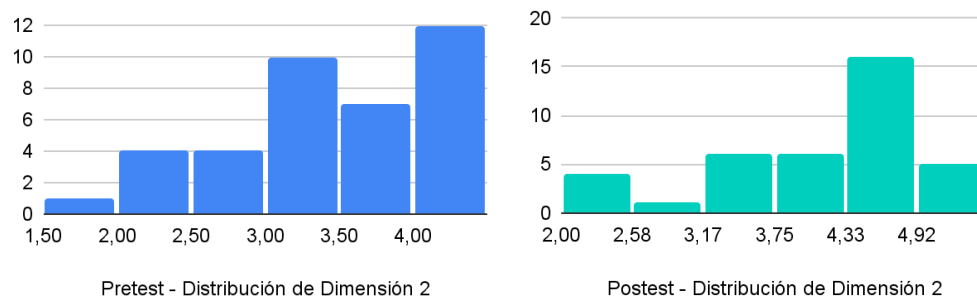
La dimensión 1 se evaluó a partir del promedio de las preguntas 1 a 4. En el pretest, el puntaje promedio fue de 3.25, lo que refleja una posición neutral de los trabajadores respecto a la recepción de expedientes judiciales, con una dispersión moderada (DE =

0.99). Los valores oscilaron entre 1.25 y 4.75, y la distribución presentó una forma aplanada y asimetría negativa. En el posttest, se evidenció una mejora significativa, con un puntaje promedio de 4.24, indicando una percepción positiva. La mediana alcanzó el valor máximo de la escala (5), y la dispersión fue ligeramente mayor (DE = 1.02), con valores entre 2 y 5. La distribución se mantuvo ligeramente aplanada y con asimetría negativa.

Dimensión 2 (D2): Elaboración y Envío de Recursos Legales

Figura 66

Histograma de la Dimensión 2



Nota. Elaboración Propia.

Tabla 21

Medidas clave de la Dimensión 2

Dimensión 2	Pretest	Posttest
Media	3,309211	4,144737
Mediana	3,375	4,5
Desv. estándar	0,7316083	0,8613934
Varianza	0,535	0,742
Asimetría	-0,615	-1,078
Curtosis	-0,667	-0,02
Rango	2,5	2,75
Mínimo	1,75	2,25
Máximo	4,25	5

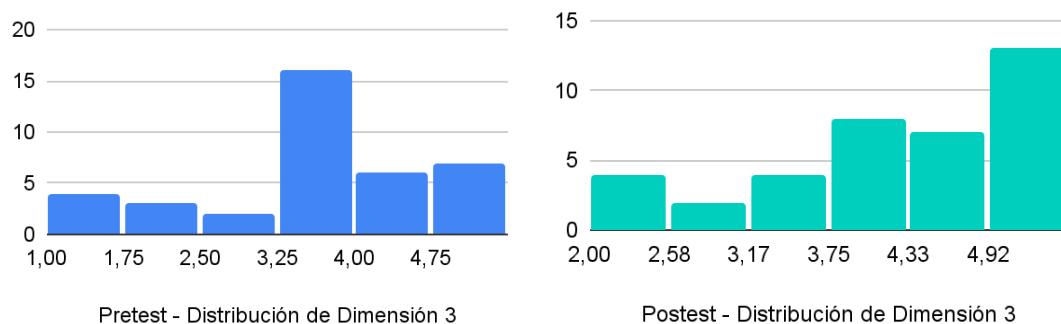
Nota. Elaboración Propia.

La dimensión 2 se evaluó mediante el promedio de las preguntas 5 a 8. En el pretest, el puntaje promedio fue de 3.30, lo que refleja una percepción neutral por parte de los trabajadores, con una dispersión moderada ($DE = 0.73$). Los valores oscilaron entre 1.75 y 4.25, y la distribución presentó una forma aplanada y asimetría negativa. En el posttest, se evidenció una mejora significativa, alcanzando un puntaje promedio de 4.14, lo que indica una percepción positiva. La mediana se situó en 4.5, cercana al valor máximo de la escala. La dispersión fue ligeramente mayor ($DE = 0.86$), con valores entre 2.25 y 5. Asimismo, la distribución se mantuvo ligeramente platicúrtica y con asimetría negativa.

Dimensión 3 (D3): Planificación y Participación en Audiencias

Figura 67

Histograma de la Dimensión 3



Nota. Elaboración Propia.

Tabla 22

Medidas clave de la Dimensión 3

Dimensión 3	Pretest	Posttest
Media	3,526316	4,190789
Mediana	3,75	4,5
Desv. estándar	1,1162036	0,8918441
Varianza	1,246	0,795
Asimetría	-0,645	-0,934
Curtosis	-0,323	-0,318
Rango	3,75	2,75

Mínimo	1,25	2,25
Máximo	5	5

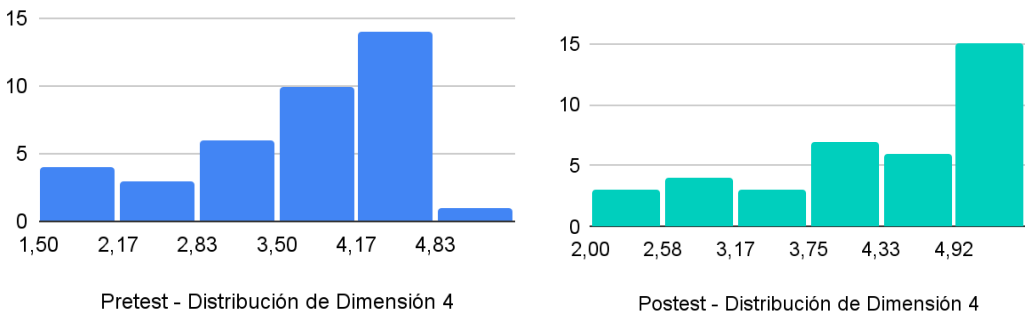
Nota. Elaboración Propia.

La dimensión 3 se evaluó mediante el promedio de las preguntas 9 a 12. En el pretest, el puntaje promedio fue de 3.52, lo que refleja una percepción neutral por parte de los trabajadores, con una alta dispersión ($DE = 1.11$). Los valores oscilaron entre 1.25 y 5, y la distribución presentó una forma aplanada y asimetría negativa. En el posttest, se evidenció una mejora significativa, alcanzando un puntaje promedio de 4.19, lo que indica una percepción positiva. La mediana se situó en 4.5, cercana al valor máximo de la escala. La dispersión disminuyó ($DE = 0.89$), con valores entre 2.25 y 5. Asimismo, la distribución se mantuvo platicúrtica y con asimetría negativa.

Dimensión 4 (D4): Revisión y Seguimiento de Expedientes

Figura 68

Histograma de la Dimensión 4



Nota. Elaboración Propia.

Tabla 23

Medidas clave de la Dimensión 4

Dimensión 4	Pretest	Posttest
Media	3,697368	4,210526
Mediana	3,875	4,5
Desv. estándar	0,9464691	0,8939102

Varianza	0,896	0,799
Asimetría	-0,736	-0,852
Curtosis	-0,448	-0,632
Rango	3,25	2,75
Mínimo	1,75	2,25
Máximo	5	5

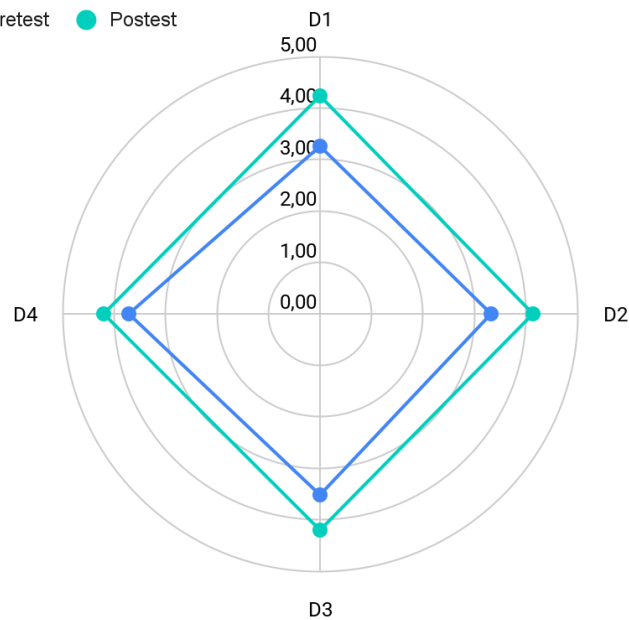
Nota. Elaboración Propia.

La dimensión 4 se evaluó mediante el promedio de las preguntas 13 a 16. El pretest mostró un puntaje promedio de 3.70, lo que refleja una percepción neutral por parte de los trabajadores, con una dispersión moderada ($DE = 0.94$). Los valores oscilaron entre 1.75 y 5, y la distribución presentó una forma aplanada y asimetría negativa. En el posttest, se evidenció una mejora significativa, alcanzando un puntaje promedio de 4.21, lo que indica una percepción positiva. La mediana se situó en 4.5, cercana al valor máximo de la escala. La dispersión disminuyó ligeramente ($DE = 0.89$), con valores entre 2.25 y 5. Asimismo, la distribución se mantuvo platicúrtica y con asimetría negativa.

Comparación de Dimensiones. Como se puede apreciar en la Figura 69, se tienen las puntuaciones obtenidas para las 4 dimensiones tanto en el pretest como en el posttest. Con respecto a los datos del pretest, se observa que D1 y D2 obtuvieron las puntuaciones más bajas cercanas a 3. En contraste con la dimensión de D4 que obtuvo el puntaje más alto en el pretest con un valor cercano a 4. Por otra parte, en el posttest se logró una gran mejora en todas las dimensiones, siendo la más notable la de D1 que obtuvo un incremento de casi un punto.

Figura 69

Comparación de las dimensiones de Gestión de Expedientes Judiciales



Nota. Elaboración Propia.

6.1.3. Análisis Estadístico Inferencial

6.1.3.1. Análisis de Normalidad. Antes de poder proceder con una prueba paramétrica, se debe corroborar que los datos siguen una distribución normal. Tal como lo indica Field (2013), para aplicar una prueba t de muestras relacionadas primero hay que verificar que las diferencias de los puntajes obtenidos en el pretest y posttest se aproximen a una distribución normal. Para ello se utilizará la prueba de Shapiro-Wilk dado a que $n < 50$. Los resultados obtenidos son los siguientes:

Tabla 24

Prueba de Shapiro-Wilk

	Shapiro-Wilk			Normalidad ($p > 0.05$)
	Estadístico	gl	p-valor	
Variable dependiente	0.981	38	0.758	Sí

Dimensión 1 (D1)	0.945	38	0.063	Sí
Dimensión 2 (D2)	0.923	38	0.012	No
Dimensión 3 (D3)	0.965	38	0.265	Sí
Dimensión 4 (D4)	0.932	38	0.024	No

Nota. Elaboración Propia.

Como se observa, el valor de significancia (p) es mayor que 0.05 para la variable dependiente, la dimensión 1 y la dimensión 3, lo que indica que estas presentan una distribución normal. En consecuencia, para la contrastación de hipótesis se aplicará la prueba paramétrica t de muestras relacionadas en dichos casos. Por otro lado, en las dimensiones 2 y 4, el valor de p es menor o igual a 0.05, evidenciando la ausencia de normalidad; por ello, se utilizará la prueba no paramétrica de Wilcoxon.

6.1.3.2. Contrastación de hipótesis. A continuación, se muestran las pruebas de hipótesis tanto para la hipótesis general y específicas.

Hipótesis General

- Hipótesis Nula: La reingeniería del sistema web SIE no mejora la gestión de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.
- Hipótesis Alternativa: La reingeniería del sistema web SIE mejora la gestión de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.

Tabla 25

Prueba t para la hipótesis general

Diferencias emparejadas				Significación	
Media	Correlación	t	gl	p valor de un factor	p valor de dos factores
-0.7516	0.92	-12.719	37	<.001	<.001

Nota. Elaboración Propia.

Como se muestra en la Tabla 25, el valor de significancia obtenido en la prueba t es menor a 0.05, por lo que se acepta la hipótesis alternativa, permitiendo afirmar que la

reingeniería del sistema web SIE mejora la gestión de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.

Hipótesis Específica 1

- Hipótesis Nula: La reingeniería del sistema web SIE no mejora la recepción y consulta de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.
- Hipótesis Alternativa: La reingeniería del sistema web SIE mejora la recepción y consulta de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.

Tabla 26

Prueba t para la hipótesis específica 1

Diferencias emparejadas				Significación	
Media	Correlación	t	gl	p valor de un factor	p valor de dos factores
-0.9934	0.857	-11.329	37	<.001	<.001

Nota. Elaboración Propia.

Como se muestra en la Tabla 26, el valor de significancia obtenido en la prueba t es menor a 0.05, por lo que se acepta la hipótesis alternativa, permitiendo afirmar que la reingeniería del sistema web SIE mejora la recepción y consulta de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.

Hipótesis Específica 2

- Hipótesis Nula: La reingeniería del sistema web SIE no mejora la elaboración y envío de recursos legales de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.

- Hipótesis Alterna: La reingeniería del sistema web SIE mejora la elaboración y envío de recursos legales de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.

Tabla 27

Prueba de wilcoxon para la hipótesis específica 2

Wilcoxon

Z	Sig. asin (lateral)
-5.392	<0.001

Nota. Elaboración Propia.

Como se muestra en la Tabla 27, el valor de significancia obtenido en la prueba de Wilcoxon es menor a 0.05, por lo que se acepta la hipótesis alterna, permitiendo afirmar que la reingeniería del sistema web SIE mejora la elaboración y envío de recursos legales de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.

Hipótesis Específica 3

- Hipótesis Nula: La reingeniería del sistema web SIE no mejora la planificación y participación en audiencias de los miembros de la Procuraduría Pública del Sector Interior - MININTER.
- Hipótesis Alterna: La reingeniería del sistema web SIE mejora la planificación y participación en audiencias de los miembros de la Procuraduría Pública del Sector Interior - MININTER.

Tabla 28*Prueba t para la hipótesis específica 3*

Diferencias emparejadas				Significación	
Media	Correlación	t	gl	p valor de un factor	p valor de dos factores
-0.66447	0.879	-7.587	37	<.001	<.001

Nota. Elaboración Propia.

Como se muestra en la Tabla 28, el valor de significancia obtenido en la prueba t es menor a 0.05, por lo que se acepta la hipótesis alterna, permitiendo afirmar que la reingeniería del sistema web SIE mejora la planificación y participación en audiencias de los miembros de la Procuraduría Pública del Sector Interior - MININTER.

Hipótesis Específica 4

- Hipótesis Nula: La reingeniería del sistema web SIE no mejora la revisión y seguimiento de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.
- Hipótesis Alterna: La reingeniería del sistema web SIE mejora la revisión y seguimiento de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.

Tabla 29*Prueba de wilcoxon para la hipótesis específica 4*

Wilcoxon	
Z	Sig. asin (lateral)

-4.836	<0.001
--------	--------

Nota. Elaboración Propia.

Como se muestra en la Tabla 29, el valor de significancia obtenido en la prueba de Wilcoxon es menor a 0.05, por lo que se acepta la hipótesis alterna, permitiendo afirmar que la reingeniería del sistema web SIE mejora la revisión y seguimiento de expedientes judiciales en la Procuraduría Pública del Sector Interior - MININTER.

6.1.3.3. Análisis del tamaño del efecto

Tabla 30

d de cohen

	Standardizer	Estimación de puntos	Intervalo de confianza al 95%	
			Inferior	Superior
d de cohen	0.3642858	2.063	1.493	2.625
Corrección de Hedges	0.3718839	2.021	1.462	2.571

Nota. Elaboración Propia.

En la Tabla 30 se verifica cual es el tamaño del efecto a través de la d de Cohen para la hipótesis general. Una d de Cohen mayor de 0.2 indica un efecto débil; un valor de 0.5, un efecto moderado; y un valor mayor a 0.8, indica un efecto grande. En este proyecto se obtuvo una d de Cohen de 2.063 lo que demuestra que el sistema SIE tiene un impacto muy significativo.

6.2. Análisis de Resultados para la población 2

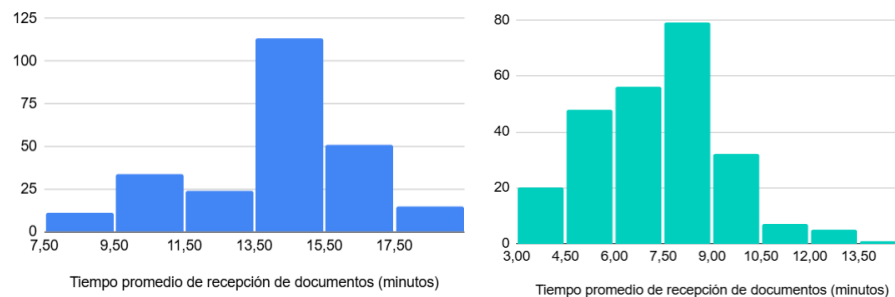
6.2.1. Análisis Estadístico Descriptivo

La ficha de registro utilizada para la recolección de datos de la población 2 consta de 2 indicadores de tiempo I1 e I2. El detalle de los resultados obtenidos se muestra a continuación:

Indicador 1 (I1): Tiempo de recepción de documentos

Figura 70

Distribución del indicador I1 en el grupo de control y en el grupo experimental



Nota. Elaboración Propia.

Tabla 31

Medidas descriptivas del indicador I1 en el grupo de control y en el grupo experimental

I1	Grupo de control	Grupo experimental
Media	14.0915	7.3776
Mediana	14.3500	7.5000
Desv. estándar	1.59848	2.01834
Varianza	2.55517	4.074
Asimetría	-.513	0.463
Curtosis	.155	0.710
Rango	11.70	11.44
Mínimo	7.55	3.34
Máximo	19.25	14.78

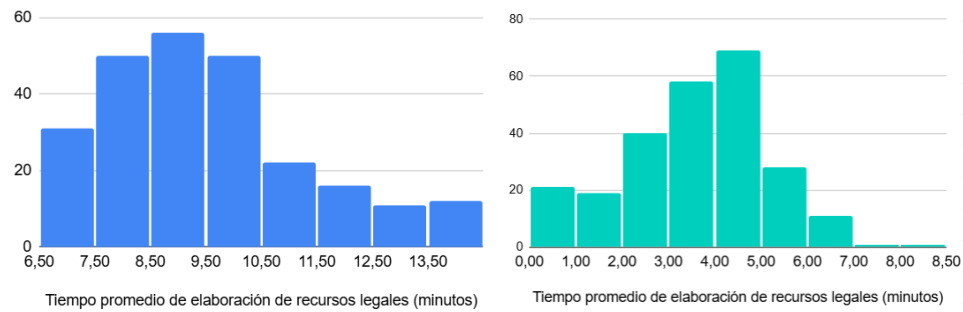
Nota. Elaboración Propia.

Los resultados evidencian una diferencia significativa entre el grupo de control y el grupo experimental en el indicador I1 (véase Tabla 31 y Figura 70). En el grupo de control, el tiempo promedio se situó alrededor de los 14 minutos, con una concentración en el intervalo de 13.5 a 15.5 minutos. En contraste, el grupo experimental registró un promedio aproximado de 7.4 minutos, concentrándose principalmente entre 6 y 9 minutos.

Indicador 2 (I2): Tiempo promedio de elaboración de recursos legales

Figura 71

Distribución del indicador I2 en el grupo de control y grupo experimental



Nota. Elaboración Propia.

Tabla 32

Medidas descriptivas del indicador I2 en el grupo de control y grupo experimental

I2	Grupo de control	Grupo experimental
Media	9.5610	3.59608
Mediana	9.3050	3.83350
Desv. estándar	1.86730	1.577729
Varianza	3.487	2.489
Asimetría	0.679	-0.253
Curtosis	-0.172	-0.258
Rango	7.70	8.038
Mínimo	6.50	0.006
Máximo	14.20	8.044

Nota. Elaboración Propia.

Los resultados evidencian una diferencia significativa entre el grupo de control y el grupo experimental en el indicador I2 (véase Tabla 32 y Figura 71). En el grupo de control, el tiempo promedio se ubicó entre 9 y 10 minutos, con una mayor concentración en el intervalo de 8 a 10 minutos, lo que refleja un proceso relativamente uniforme, aunque con ligera asimetría debido a algunos valores elevados. En contraste, el grupo experimental

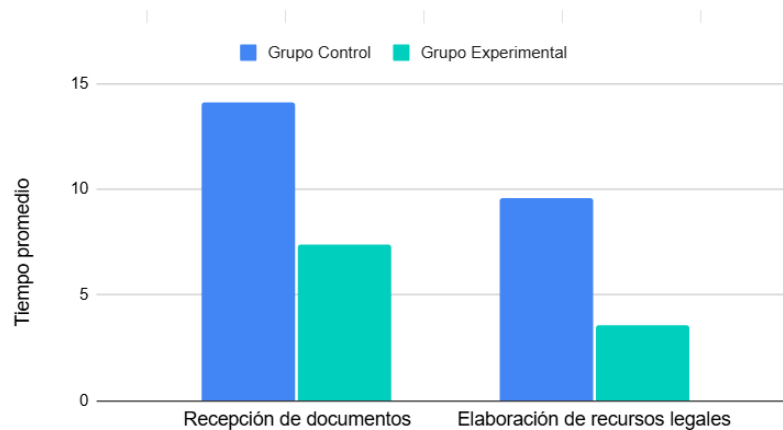
registró un promedio cercano a 3.6 minutos, concentrándose principalmente entre 3 y 5 minutos, lo que evidencia una notable mejora en la agilidad del proceso.

Cabe señalar que algunos valores muy reducidos observados en el grupo experimental corresponden a recursos legales del tipo “formato de archivo”, los cuales son completamente estandarizados y no requieren intervención manual del abogado, por lo que su tiempo de elaboración equivale únicamente al tiempo que el sistema tarda en generarlos y descargarlos.

Comparación de indicadores

Figura 72

Comparación de promedios para I1 y I2 entre el grupo de control y el grupo experimental



Nota. Elaboración Propia.

La comparación de los promedios de los indicadores I1 e I2 entre el grupo de control y el grupo experimental, evidencian el impacto positivo de la mejora implementada en el sistema SIE (véase Figura 72). En el indicador I1, el tiempo de recepción de documentos se redujo aproximadamente de 14 a 7.38 minutos, mientras que en el indicador I2, el tiempo de elaboración de recursos legales disminuyó de cerca de 9.5 a 3.6 minutos.

6.2.2. *Análisis Estadístico Inferencial*

6.2.2.1. Prueba de normalidad. Antes de seleccionar la prueba estadística apropiada (paramétrica o no paramétrica), es necesario verificar que los datos analizados presentan una distribución normal. De acuerdo con Hernández-Sampieri et al. (2014), para la aplicación de la prueba t de muestras relacionadas se requiere que las diferencias entre los puntajes obtenidos se aproximen a una distribución normal. En ese sentido, con la finalidad de comprobar dicho supuesto, se emplea la prueba de Kolmogorov-Smirnov, considerando que el tamaño de la muestra es mayor a 50. Los resultados obtenidos son los siguientes:

Tabla 33

Prueba de Kolmogorov-Smirnov para Tiempo de Recepción

	N	Sig. Asin. Bilateral	Sig. Asin. Monte Carlo
G. Control	248	<.001	<.001
G. Experimental	248	.001	.001

Nota. Elaboración Propia.

Los resultados de la prueba de normalidad muestran que, tanto en el grupo de control como en el grupo experimental para métrica de Tiempo de Recepción, el valor de significancia es menor a 0.05 (Sig. < 0.05). En el grupo de control, con una muestra de 248 observaciones, se obtuvo un valor de significancia menor a 0.001, mientras que en el grupo experimental el valor fue de 0.001.

Tabla 34

Prueba de Kolmogorov-Smirnov para Tiempo de Elaboración

	N	Sig. Asin. Bilateral	Sig. Asin. Monte Carlo
G. Control	248	<.001	<.001
G. Experimental	248	0.009	0.011

Nota. Elaboración Propia.

De la misma manera, los resultados de la prueba de normalidad para métrica de Tiempo de Elaboración evidencian que el valor de significancia es menor a 0.05 (Sig. < 0.05). En el grupo de control, con una muestra de 248 observaciones, se obtuvo un valor de significancia menor a 0.001, mientras que en el grupo experimental el valor fue de 0.009.

Estos resultados indican que los datos no siguen una distribución normal en ninguno de los dos grupos evaluados. Por lo tanto, para el análisis comparativo entre el grupo de control y el grupo experimental, resulta más apropiado el uso de pruebas estadísticas no paramétricas, las cuales no requieren que los datos se distribuyan normalmente y se ajustan mejor a las características de la información recolectada.

6.2.2.2. Prueba de U de Mann-Whitney. Se aplicó la prueba U de Mann-Whitney para el indicador *I1*, obteniéndose un valor de $U = 1629.00$, $Z = -18.247$ y una significancia bilateral $p < .001$, lo cual evidencia diferencias estadísticamente significativas entre el grupo de control y el grupo experimental para el indicador Tiempo de recepción de documentos (*I1*).

Tabla 35

Prueba de U de Mann-Whitney para I1

U de Mann-Withney	1629,00
Z	-18.247
Sig. asin. (bilateral)	<.001

Nota. Elaboración Propia.

Por otro lado, también, se aplicó la prueba U de Mann-Whitney para el indicador *I2*, obteniéndose un valor de $U = 102,00$, $Z = -19.204$ y una significancia bilateral $p < .001$, lo cual evidencia diferencias estadísticamente significativas entre el grupo de control y el grupo experimental para el indicador *Tiempo promedio de elaboración de recursos legales* (*I2*).

Tabla 36*Prueba de U de Mann-Whitney para I2*

U de Mann-Withney	102,00
Z	-19.204
Sig. asin. (bilateral)	<.001

Nota. Elaboración Propia.

6.3. Resumen de resultados obtenidos

A continuación, se muestra los resultados obtenidos por módulo:

Tabla 37

Resumen de resultados obtenidos

Código	Necesidad	User Stories	Módulo	Indicador	Instrumento	Resultado Indicador
N01	Visualización y búsqueda de legajos	SIE-11 a SIE-18	Consulta Legajo	Nivel de percepción sobre la completitud, precisión y accesibilidad de la documentación de los expedientes recibidos.	Cuestionario (Ítem 1,2,3)	4.24/5 (Casi siempre)
N02	Gestión documental del legajo	SIE-19 a SIE-25	Consulta Legajo	Tiempo promedio de elaboración de recursos legales.	Ficha de registro	7,38 minutos
				Nivel de percepción sobre la rapidez en la elaboración de recursos legales	Cuestionario (Ítem 5,6)	4.07/5 (Casi siempre)
N03	Control del estado y exportación	SIE-25 a SIE-29	Consulta Legajo	Nivel de aceptación de recursos legales por otras entidades	Cuestionario (Ítem 7,8)	4.21/5 (Casi siempre)

N04	Creación y actualización de legajos	SIE-04 a SIE-10	Recepción Legajo	Tiempo promedio de recepción de expedientes.	Ficha de registro	3.59 minutos
N05	Extracción rápida de información contenida en los documentos	SIE-18 a SIE-21	Recepción Legajo			
N06	Validación y corrección de datos	SIE-32 a SIE-38	Recepción Legajo			
N07	Programación de audiencias	SIE-41 a SIE-42	Audiencias	Nivel de percepción sobre el cumplimiento de los plazos en la programación de audiencias.	Cuestionario (Item 9)	4.21/5 (Casi siempre)
N08	Modificación y reprogramación	SIE-43 a SIE-47	Audiencias	Nivel de asistencia a audiencias de los procuradores adjuntos.	Cuestionario (Item 10, 11,12)	4.18/5 (Casi siempre)
				Grado de identificación de tendencias o patrones recurrentes en los expedientes.	Cuestionario (Item 13)	3.94/5 (Algunas veces sí, algunas veces no)
N09	Visualización de indicadores	SIE-39 a SIE-41	Reportes	Grado de seguimiento del estado de los expedientes.	Cuestionario (Item 14, 15,16)	4.29/5 (Casi siempre)

Nota. Elaboración Propia.

6.4. Discusión de Resultados

Al contrastar los resultados del presente estudio con investigaciones previas, se evidencia coherencia con los hallazgos reportados por Acoba et al. (2020), quienes analizan la incorporación de sistemas tecnológicos en procesos administrativos del ámbito judicial. En el contexto específico de la Procuraduría Pública del Ministerio del Interior, los resultados muestran una mejora significativa en los tiempos operativos asociados a la gestión de expedientes judiciales, particularmente en la recepción de documentos y la elaboración de recursos legales. Estos hallazgos coinciden con lo señalado por Rojas et al. (2022), quienes destacan que la reducción de tiempos constituye un factor clave para optimizar la atención de casos y mejorar la eficiencia operativa en instituciones vinculadas a la administración de justicia. Asimismo, la simplificación de tareas legales observada en el presente estudio se alinea con lo expuesto por Zhong et al. (2020), quienes sostienen que la optimización de procesos documentales contribuye a una gestión más ordenada y eficiente de la información legal.

En relación con la estandarización de la implementación, los resultados obtenidos evidencian que el uso de inteligencia artificial generativa permitió estructurar y procesar información documental de manera consistente dentro del contexto institucional analizado. Estos resultados guardan relación con los estudios de Kainan y Xin (2021) y Anu y Sangeetha (2021), quienes reportan mejoras significativas al aplicar modelos de aprendizaje automático sobre documentos previamente estandarizados. Sin embargo, a diferencia de dichos estudios, en el presente trabajo la información procesada no sigue un único formato rígido, lo que exigió un enfoque más flexible en el diseño y ajuste de la solución tecnológica implementada.

Cabe precisar que los fundamentos conceptuales sobre el uso de modelos avanzados para el procesamiento de documentos no estructurados y altamente variables fueron abordados en la base teórica. En coherencia con lo señalado por Gu et al. (2020), cuando se trabaja con información documental heterogénea resulta necesario emplear modelos de mayor complejidad. Mientras que en su estudio se utilizó el modelo BERT, en la presente investigación se optó por un modelo de lenguaje de gran escala (LLM), específicamente LLAMA, considerando la evolución tecnológica actual. La aplicación de este modelo permitió mejorar la extracción de información relevante y

reducir los tiempos percibidos en los procesos analizados, lo cual se reflejó en los resultados del grupo experimental.

Finalmente, es importante señalar que, al interpretar los resultados obtenidos, es necesario considerar la posible influencia del efecto Hawthorne, el cual se refiere a los cambios en el comportamiento de los participantes derivados del conocimiento de que están siendo observados o de que forman parte de una intervención específica. En el presente estudio, los miembros del grupo experimental eran conscientes de la implementación de una mejora tecnológica en los procesos de gestión de expedientes judiciales, lo que pudo haber generado una mayor predisposición al uso adecuado del sistema y a una ejecución más cuidadosa de las actividades evaluadas respecto a la población. No obstante, esta consideración no invalida los resultados obtenidos, sino que permite interpretarlos con cautela, reconociendo que parte de la mejora percibida puede estar influenciada por factores conductuales propios de contextos de implementación.

Conclusiones

- Se comprobó, con un nivel de confianza del 99%, que la reingeniería del sistema web SIE mejoró significativamente la percepción de los usuarios sobre la gestión de expedientes judiciales en la Procuraduría Pública del Sector Interior – MININTER. Esta mejora en la percepción del desempeño estuvo acompañada por un tamaño del efecto grande (d de Cohen = 2.06), evidenciando un cambio relevante en la valoración global del sistema, cuyo puntaje promedio pasó de 3.45 en el pretest a 4.20 en el posttest. Asimismo, se registraron mejoras significativas en todas las dimensiones evaluadas, destacando la dimensión de Recepción de Expedientes Judiciales, que presentó el mayor incremento en la escala de Likert.
- En relación con la recepción de expedientes judiciales, la implementación del módulo desarrollado como parte de la reingeniería permitió optimizar el proceso de registro de la información de los documentos de entrada mediante el uso del modelo de lenguaje Llama 3.1. Este cambio no solo impactó en la percepción de los usuarios, sino que también se reflejó en los tiempos observados: mientras que en el grupo de control el tiempo promedio de recepción de documentos fue de aproximadamente 14.09 minutos, en el grupo experimental dicho tiempo se redujo a un promedio de 7.38 minutos. Esta disminución respalda los resultados perceptuales, evidenciando que las nuevas funcionalidades contribuyeron a una mayor agilidad operativa en Mesa de Partes.
- Respecto a la elaboración y envío de recursos legales, se logró disminuir el tiempo de creación de los cuatro principales recursos utilizados por la Procuraduría, permitiendo su generación a partir de la información almacenada en el sistema SIE y su envío directo por correo electrónico. En el grupo de control, el tiempo promedio de elaboración de recursos legales fue de 9.56 minutos, mientras que en el grupo experimental este tiempo se redujo a 3.60 minutos en promedio. Esta reducción sustenta la mejora percibida por los usuarios, quienes identificaron una disminución del esfuerzo y del tiempo requerido para esta actividad.
- En cuanto a la planificación y participación en audiencias, así como a la revisión y seguimiento de expedientes, se implementaron módulos que no sustituyen el criterio legal

del abogado o del procurador, pero que facilitan el acceso, la organización y la visualización de la información relevante. Estas funcionalidades permitieron a los usuarios contar con una visión general del estado de los expedientes, de las audiencias programadas y de las tendencias jurídicas, lo que se tradujo en una mejor percepción del control y seguimiento de los casos, apoyada por una reducción en los tiempos de consulta y análisis de la información.

Recomendaciones

- A partir de los resultados obtenidos en la presente investigación, se recomienda que futuros estudios analicen la aplicación del sistema SIE en otras procuradurías públicas con funciones similares, con el objetivo de contrastar los hallazgos en distintos contextos institucionales. Esta recomendación no busca generalizar los resultados del presente estudio, sino ampliar la evidencia empírica sobre el impacto de soluciones tecnológicas en la gestión de expedientes judiciales, considerando las particularidades operativas de cada entidad.
- Asimismo, se recomienda que en investigaciones posteriores se incorpore un periodo de seguimiento prolongado posterior a la implementación del sistema, con la finalidad de evaluar la estabilidad de los resultados en el tiempo. Este monitoreo permitirá reducir la influencia de factores conductuales temporales, como el efecto Hawthorne, y determinar si las mejoras observadas en los indicadores se mantienen una vez que los usuarios se han adaptado plenamente al uso cotidiano del sistema.
- Finalmente, se sugiere explorar, en futuros trabajos de investigación, la adaptación del módulo de recepción de expedientes a otras ramas del Derecho, como el derecho civil o tributario, utilizando conjuntos de datos específicos de cada especialidad. Esta línea de investigación permitiría evaluar la viabilidad del enfoque propuesto bajo distintos marcos normativos y documentales, manteniendo el uso de técnicas de inteligencia artificial generativa como apoyo a la gestión documental, pero ajustadas a cada dominio jurídico.

Limitaciones

- La presente investigación presenta algunas limitaciones que deben ser consideradas al interpretar los resultados. En primer lugar, la solución propuesta no puede generalizarse automáticamente a todas las procuradurías ni a otros organismos del sistema judicial, dado que el modelo fue entrenado y ajustado con una gama específica de delitos y documentos propios del ámbito penal. Por tanto, su desempeño está directamente condicionado por el tipo de información utilizada durante su configuración.
- En segundo lugar, cualquier cambio en la lógica de negocio, en los procedimientos internos o en la estructura documental de la institución implicará la actualización de los prompts utilizados por el modelo de lenguaje, lo que requiere mantenimiento técnico y revisión periódica para asegurar la coherencia de los resultados generados por la IA.
- Finalmente, como se señaló en la interpretación de resultados, no se puede descartar completamente la influencia del efecto Hawthorne, dado que los participantes del grupo experimental eran conscientes de la implementación de una mejora tecnológica. Esta situación pudo haber influido parcialmente en su comportamiento durante la etapa de evaluación, constituyéndose en una limitación propia de estudios aplicados en entornos organizacionales reales.

Referencias Bibliográficas

- Abeliuk, A. & Gutierrez, C. (2021). Historia y evolución de la inteligencia artificial. *Inteligencia Artificial*, 14-21
<https://revistasdex.uchile.cl/index.php/bits/article/download/2767/2700>
- Acoba, A. , Cunanan, C., Merencilla, N., Tejada, R., & Ligayo, M. (2020). *JuCo-IS: A Development of Web-Based Information System in Judicial Regional Trial Court*. 2020 IEEE 10th International Conference on System Engineering and Technology (ICSET). 10.1109/icset51301.2020.9265362
- Adhikary, S., Sen, P., Roy, D., & Ghosh, K. (2024). *A case study for automated attribute extraction from legal documents using large language models*. *Artificial Intelligence and Law*. <https://doi.org/10.1007/s10506-024-09425-7>
- Aguiló, J. (2019). En defensa del Estado constitucional de Derecho. *Doxa. Cuadernos de Filosofía del Derecho*, 42, 85-100. <https://doi.org/10.14198/DOXA2019.42.04>
- Amatriain, X. (2024). *Prompt design and engineering: Introduction and advanced methods*. arXiv. <https://doi.org/10.48550/arXiv.2401.14423>
- AMAZON. (2020). ¿Qué es el ajuste de hiper parámetros?. Amazon Web Services. <https://aws.amazon.com/es/what-is/hyperparameter-tuning/>
- Anu, T., & Sangeetha, S. (2021). Semi-supervised, knowledge-integrated pattern learning approach for fact extraction from judicial text. *Expert Systems*, 38(3). <https://doi.org/10.1111/exsy.12656>
- Ateia, S., Kruschwitz, U., Scholz, M., Koschmider, A., & Almohaishi, M. (2025). *LLM-based information extraction to support scientific literature research and publication workflows*. arXiv. <https://doi.org/10.48550/arXiv.2510.04749>
- Ávila-Tomás, J., Mayer-Pujadas, M. & Quesada-Varela, V. (2020) La inteligencia artificial y sus aplicaciones en medicina I: introducción antecedentes a la IA y robótica. *Atención Primaria*, 52(10), 778-784, <https://doi.org/10.1016/j.aprim.2020.04.013>.
- Bosco-Hernández, H. (03 de Agosto de 2023).Avantius: el sistema judicial de expedientes digitales con el que Navarra es pionera. Ser. <https://cadenaser.com/navarra/2023/08/03/avantius-el-sistema-judicial-de-expedientes-digitales-con-el-que-navarra-es-pionera-radio-pamplona/>
- Brown, M. & Hale, K. (2014). *Applied Research Methods in Public and Nonprofit Organizations*. Jossey-Bass. <https://www.wiley.com/en-us/Applied+Research+Methods+in+Public+and+Nonprofit+Organizations-p-9781118737361>
- Castañeda, W. (2022). Sistema de Información Web, Utilizando Metodología Scrum para la Gestión de Expedientes Judiciales en la Procuraduría Pública Regional de San

- Martín [Tesis de Grado, Universidad Cesar Vallejo]. Repositorio Digital Institucional de la Universidad Cesar Vallejo.
[https://repositorio.ucv.edu.pe/bitstream/handle/20.500.12692/95704/Casta%
 eda_TWA-SD.pdf?sequence=8&isAllowed=y](https://repositorio.ucv.edu.pe/bitstream/handle/20.500.12692/95704/Casta%c3%b1eda_TWA-SD.pdf?sequence=8&isAllowed=y)
- Chen, H., & Luo, X. (2019). An automatic literature knowledge graph and reasoning network modeling framework based on ontology and natural language processing. *Advanced Engineering Informatics*, 42, 100959.
<https://doi.org/10.1016/j.aei.2019.100959>
- Coll, A., & Restrepo, R. (2021). Approach towards a structural understanding of e-justice. *Cinta de moebio*, (70), 81-93. <https://dx.doi.org/10.4067/s0717-554x2021000100081>
- Congreso de la República del Perú. (2004). Código Procesal Penal (Decreto Legislativo N.º 957) <https://diariooficial.elperuano.pe/Normas/obtenerDocumento?idNorma=70003>
- Cordella, A. & Contini, F. (2020). Digital Technologies for Better Justice. <https://publications.iadb.org/publications/english/document/DigitalTechnologies-for-Better-Justice-A-Toolkit-for-Action.p>
- COURT SERVICES. (12 de septiembre de 2022). The court system.Court Services. <https://courts.vic.gov.au/court-system>
- Dong, J. (2023). Natural Language Processing Pretraining Language Model for Computer Intelligent Recognition Technology. *ACM Trans.* <https://doi.org/10.1145/3605210>
- Erazo-Luzuriaga, A. F., Ramos-Secaira, F. M., Galarza-Sánchez, P. C., & Boné-Andrade, M. F. (2023). La inteligencia artificial aplicada a la optimización de programas informáticos. *Journal of Economic and Social Science Research*, 3(1), 48–63.
<https://doi.org/10.55813/gaea/jessr/v3/n1/61>
- Esquicha, J. (2022). Implementación de una Plataforma Web de Almacenamiento y difusión de planos arquitectónicos antiguos usando OCR y tecnologías web. [Tesis para optar el grado de bachiller]. Universidad Catolica de Santa Maria
- Estevez, E., Linares, S. & Fillottrani, P. (2020) Prometea. Transformando la administración de justicia con herramientas de inteligencia artificial. Banco Interamericano de Desarrollo (BID). <https://publications.iadb.org/es/prometeatransformando-la-administracion-de-justiciacon-herramientas-de-inteligencia-artificia>
- Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics* (4th ed.).
- Gadekallu, T., Huynh-The, T., Wang, W., Yenduri, G., Ranaweera, P. & Liyanage, M. (2022). Blockchain for the Metaverse: A Review. Cornell University. <https://arxiv.org/abs/2203.09738>

- García, E. (2023) Derecho Penal contemporáneo. *Revista Ciencia Multidisciplinaria CUNORI*, 7(1), 119–129. <https://doi.org/10.36314/cunori.v7i1.212>
- GeeksforGeeks. (2025, 15 de noviembre). *GenAI Ops*.
<https://www.geeksforgeeks.org/artificial-intelligence/genaiops/>
- GOBIERNO DE CANARIAS. (2022). Desarrollo del Sistema de Gestión Judicial Atlante web y su evolución. Gobierno de Canarias.
<https://www.gobiernodecanarias.org/justicia/Formacion/FormacionAtlante/FormacionDesarrollo/>
- Gu, L., Zhang, W., Wang, Y., Li, B., & Mao, S. (2020). Named Entity Recognition in the Judicial Field Based on BERT-BiLSTM-CRF Model. In 2020 International Workshop on Electronic Communication and Artificial Intelligence (IWECAI), pp. 170-174. 10.1109/IWECAI50956.2020.00041.
- Hernández-Sampieri, R., Fernández-Collado, C. & Baptista-Lucio, P. (2014). Definición del alcance de la investigación que se realizará: exploratorio, descriptivo, correlacional o explicativo. En Hernández-Sampieri, R., Fernández-Collado, C. y Baptista-Lucio, P. *Metodología de la Investigación* (6 ed., págs. 88-101). México: McGraw-Hill.
- Huyler, D., & McGill, C. (2019, July). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches, by John Creswell and J. David Creswell. *New Horizons in Adult Education and Human Resource Development*, 31, 75-77.
<https://doi.org/10.1002/nha3.20258>
- IBM. (7 de junio de 2022). Formato JSON (JavaScript Object Notation). IBM.
<https://www.ibm.com/docs/es/baw/20.x?topic=formats-javascript-object-notation-json-format>
- Janampa, H., Karel, S. & Prado, L. (2020). Modelo de Desarrollo de Software de la Programación Extrema sobre Scrum para Gestión de Software Ágil. *Iberian Journal of Information Systems and Technologies*.
https://www.researchgate.net/profile/Hubner-Janampa/publication/349704822_Software_Development_Model_of_the_Extreme_Programming_on_Scrum_for_Agile_Software_Management/links/603daf88a6fdcc9c78081b60/Software-Development-Model-of-the-Extreme-Programming-on-Scrum-for-Agile-Software-Management.pdf
- Kalai, A. T., Nachum, O., Vempala, S. S., & Zhang, E. (2025). *Why language models hallucinate*. arXiv. <https://doi.org/10.48550/arXiv.2509.04664>
- Kainan, J., & Xin, L. (2021). Extraction Method of Judicial Language Entities Based On Regular Expression. 2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP). 10.1109/icsp51882.2021.9408748

- Khete, T., & Bakshi, A. (2022). Autonomous Assistance System for Visually Impaired using Tesseract OCR & gTTS. *Journal of Physics: Conference Series*, 2327(1), 012065. <https://dx.doi.org/10.1088/1742-6596/2327/1/012065>
- Lemontech. (6 de octubre de 2021). Control y gestión judicial a través de un sistema online. LemonTech Blog. <https://blog.lemontech.com/control-y-gestion-judicial/>
- Lim, F. J., & Tan, B. S. (2025). Overview of software re-engineering concepts, models, and approaches. *International Journal on Informatics Visualization*, 9(1), 46–56. <http://dx.doi.org/10.62527/joiv.9.1.3034>
- Liu, J. Z., & Li, X. (2025). How do judges use large language models? Evidence from Shenzhen. *Journal of Legal Analysis*, 16(1), 235–262. <https://doi.org/10.1093/jla/laae009>
- Loayza, F. (2022). Análisis de la implementación del expediente judicial electrónico en una institución del Poder Judicial. [Tesis de Maestría, Universidad César Vallejo]. Repositorio UCV. <https://repositorio.ucv.edu.pe/handle/20.500.12692/98440>
- Mainkar, V., Katkar, J., Upade, A., & Pednekar, P. (2020). Handwritten Character Recognition to Obtain Editable Text. In 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC) (pp. 599-602). Coimbatore, India. 10.1109/ICESC48915.2020.9155786.
- Matos, J. (2022). Sistema web basado en metodología XP para mejorar la gestión de expedientes en el estudio jurídico Gamión Asociados [Tesis de titulación, Universidad Nacional del Centro del Perú]. Repositorio UNCP. <https://repositorio.uncp.edu.pe/handle/20.500.12894/8552>
- Mejía, J., Agudelo, F., & Gil, K. (2021). Marco de Trabajo para Seleccionar un Patrón Arquitectónico en el Desarrollo de Software. *Revista Ibérica De Sistemas y Tecnologías De Información*, 568-581. <https://login.unmsm.lookproxy.com/?url=https://www.proquest.com/scholarly-journals/marco-de-trabajo-para-seleccionar-un-patrón/docview/2562269745/se-2>
- Microsoft. (2025). *Información general de App Service*. Microsoft Learn. <https://learn.microsoft.com/es-es/azure/app-service/overview>
- Ming Q., Yiru Z., Tianqi M., Qingfeng W. & Fanzhu J.(2020) Convolutional-neural-network-based Multilabel Text Classification for Automatic Discrimination of Legal Documents. *Sensors & Materials*, 32(8), p. 2659-2672. <https://pdfs.semanticscholar.org/08e4/ddd3b26522b80ad5d95377d805210afdb3df.pdf>
- MINISTERIO DEL INTERIOR (2019). Texto integrado del reglamento de organización y funciones del Ministerio del Interior, p 22. <https://cdn.www.gob.pe/uploads/document/file/376470/1520-2019->

IN_Aprobacion_del_Texto_Integrado_del_Reglamento_de_Organizacion_y_Funciones__ROF__del_MININTER_.pdf

MINISTERIO DEL INTERIOR (2000). Resolución N°001-2000-IN, p7-19.

https://www.mininter.gob.pe/sites/default/files/13122009232402_mof_pp_min.pdf

MINISTERIO DEL INTERIOR (2025). Compendio Estadístico Sectorial 2024, p 210-224.

https://cdn.www.gob.pe/uploads/document/file/8045905/6753666-compendio_estadistico_sector_interior_2024-f.pdf

MINISTERIO DE JUSTICIA (2020). Infografías de procesos penales según el nuevo CPP.

<https://www.gob.pe/institucion/minjus/informes-publicaciones/1455258-infografias-de-procesos-penales-segun-el-nuevo-cpp>

MINISTERIO PÚBLICO (2024). Actuación fiscal: Equipo especial de fiscales para casos con víctimas durante las protestas sociales, p7.

https://drive.google.com/file/d/1ITmxapnlyjBpVaeixxkvOk5cHMtxa_Cm/view?usp=sharing

Minsait. (2019). Minsait agiliza hasta un 50 por ciento la revisión de expedientes judiciales gracias a la inteligencia artificial. Minsait.

<https://www.minsait.com/es/actualidad/media-room/minsait-agiliza-hasta-un-50-por-ciento-la-revision-de-expedientes-judiciales>

Munir, B. (2025). *Hallucinations in legal practice: A comparative case law analysis*.

International Journal of Law, Ethics, and Technology. SSRN.
<https://doi.org/10.2139/ssrn.5265375>

OECD (2024), OECD Justice Review of Peru: Towards Effective and Transparent Justice Institutions for Inclusive Growth, OECD Publishing, Paris,

<https://doi.org/10.1787/bb556518-en>.

Oracle. (2020). ¿Qué es una base de datos? Oracle Perú.

<https://www.oracle.com/pe/database/what-is-database/>

Oré, A. & Loza, G. (2005). La Estructura del Proceso Común en el Nuevo Código Procesal Penal Peruano, p4-15.

<https://revistas.pucp.edu.pe/index.php/derechosociedad/article/view/17025/17323>

Palomino, E., Herrera, Ó., & García, A. (2025). La inteligencia artificial generativa y la justificación en la decisión judicial. *Alegatos*, (118).

<https://alegatos.azc.uam.mx/index.php/ra/article/view/1795>

Patil, S., & Desai, D. (2023). AI Enabled Blockchain solution for the Indian Judicial System. 2023 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 2023, pp. 1-6.

10.1109/ESCI56872.2023.10100151.

- Pizarro, P. (2022). Gestión jurídica y su relación con la calidad de las sentencias judiciales en la corte suprema de justicia del Perú. [Tesis de Maestría, Centro de Altos Estudios Nacionales]. Repositorio CAEN. <https://repositorio.caen.edu.pe/handle/20.500.13097/234>
- Poder Judicial (2022). Manual de usuario SDSI-SINOE-M-00008V-2022: Sistema de Notificaciones Electrónicas. https://casillas.pj.gob.pe/doc/Manual_usuario.pdf
- Pontes, J., Bruneliere, H., Tisi, M. & Sunyé, G. (2024). *Towards an In-Context LLM-Based Approach for Automating the Definition of Model Views*. In Proceedings of the 17th ACM SIGPLAN International Conference on Software Language Engineering (SLE '24), pp. 29–42. <https://doi.org/10.1145/3687997.3695650>
- RAE. (2025). Expediente judicial. En Diccionario panhispánico del español jurídico. <https://dpej.rae.es/lema/expediente-judicial>
- Ramírez, R. (2023). Tareas automatizables en la gestión de expedientes judiciales. Empecemos por el principio: cuáles son y cómo podemos identificarlas. Consejo General Abogacía Española. <https://www.abogacia.es/publicaciones/blogs/blog-de-innovacion-legal/tareas-automatizables-en-la-gestion-de-expedientes-judiciales-empecemos-por-el-principio-cuales-son-y-como-podemos-identificarlas/>
- Raschka, S. & Mirjalili, V. (2019). Python Machine Learning Third Edition. Packt Publishing Ltd. <https://www.packtpub.com/product/python-machine-learning-third-edition/9781789955750>
- Ren, Y.; Han, J., Lin, Y., Mei, X. & Zhang, L. (2022) An Ontology-Based and Deep Learning-Driven Method for Extracting Legal Facts from Chinese Legal Texts. Electronics, 11, 1821. <https://doi.org/10.3390/electronics11121821>
- Rio Labarthe, G. (2010). La acción civil en el Nuevo Proceso Penal. Derecho PUCP, (65), 221–233. <https://doi.org/10.18800/derechopucp.201002.010>
- Rodríguez, J.M. & Barrón, M.A. (2022) Inteligencia artificial y su aplicación en los sistemas de justicia en América Latina. Temas estratégicos, No. 101. Instituto Belisario Domínguez. Senado de la República. México. <http://bibliodigitalbd.senado.gob.mx/handle/123456789/5594>
- Rojas, R. D., Torres, S. L., & Anchicoque, A. (2022). Software de gestión de información: una propuesta para contrarrestar una de las causas del problema de congestión judicial en Colombia. Dictamen Libre, (30). <https://doi.org/10.18041/2619-4244/dl.30.8822>
- Rojina, A. (2008) Compendio de Derecho Civil II. Porrúa. https://www.upg.mx/wp-content/uploads/2015/10/LIBRO-13-compendio_del_derecho_civil-II.pdf
- Rubin, K. S. (2012). Essential Scrum: A Practical Guide to the Most Popular Agile Process (1st ed.). Addison-Wesley Professional. <https://github.com/manish-old/ebooks->

2/blob/master/essential-scrum-a-practical-guide-to-the-most-popular-agile-process.9780137043293.57714.pdf

- Rumiantsau, M., Vertsel, A., Hrytsuk, I., & Ballah, I. (2024). *Beyond fine-tuning: Effective strategies for mitigating hallucinations in large language models for data analytics*. arXiv. <https://doi.org/10.48550/arXiv.2410.20024>
- Ruiz, M., Román, C., Garrido, Á., & Mena, E. (2020). uAIS: An Experience of Increasing Performance of NLP Information Extraction Tasks from Legal Documents in an Electronic Document Management System. En Proceedings of the 22nd International Conference on Enterprise Information Systems - Volume 1: ICEIS (pp. 189-196). SciTePress. <https://doi.org/10.5220/0009421201890196>
- Sánchez, S. (2021). Cloud computing: fundamentos y despliegue de un servicio en la nube [Tesis de Pregrado, Universidad Autónoma de Madrid]. Repositorio UAM. https://repositorio.uam.es/bitstream/handle/10486/697844/sanchez_prado_serpio_tf_g.pdf?sequence=1
- Sarika, N., Sirisala, N. & M. Velpuru, S. (2021). CNN based Optical Character Recognition and Applications. 2021 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 10.1109/ICICT50816.2021.9358735.
- Schwaber, K. & Sutherland, J. (2020). La Guía Definitiva de Scrum: Las Reglas del Juego. Repositorio UVM. <https://repositorio.uvm.edu.ve/items/00f89266-866c-44af-8a31-a8f3c19b8c8b>
- Shadish, W, Cook, T., & Campbell, D. (2001). Experimental and Quasi-Experimental Designs for Generalized Causal Inference. https://moodle2.units.it/pluginfile.php/132646/mod_resource/content/1/Estratto_ShadishCookCampbellExperimental2002.pdf
- Singh, S. (2018). Natural Language Processing for Information Extraction. CoRR, abs/1807.02383. <http://arxiv.org/abs/1807.02383>
- Spositto, O., Ledesma, V., Procopio, G., & Bossero J. (2020). Universidad Nacional de La Matanza, 7-11 Inteligencia Artificial aplicada al Poder Judicial. http://sedici.unlp.edu.ar/bitstream/handle/10915/103381/Documento_completo.pdf-PDFA.pdf?sequence=1&isAllowed=y
- Tonmoy, S. M. T. I., Zaman, S. M. M., Jain, V., Rani, A., Rawte, V., Chadha, A., & Das, A. (2024). *A comprehensive survey of hallucination mitigation techniques in large language models*. arXiv. <https://doi.org/10.48550/arXiv.2401.01313>
- Ueno, K., Kogo, M., Kawatsu, H., Uchiumi, Y., & Tatsubori, M. (2025). *GenAIOps for GenAI Model-Agility*. arXiv. <https://doi.org/10.48550/arXiv.2502.17440>

- Ughetta, W. & Kernighan, B. (2020). The Old Bailey and OCR: Benchmarking AWS, Azure, and GCP with 180,000 Page Images. In ACM Symposium on Document Engineering 2020 (DocEng '20), New York, USA.
<https://doi.org/10.1145/3395027.3419595>
- UNITECH. (2023). ¿Qué es Iurix Mind?. Unitech. <https://www.unitech-corp.com/iurix-mind/>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., & Polosukhin, I. (2017). Attention is all you need. In Advances in neural information processing systems (pp. 30-38).
https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf
- Verheyen, G. (2019). Scrum: A smart travel companion. Van Haren (2º edición).
<https://ecole-management.com/BOOKS/1.pdf>
- Waseem, M., Liang, P., Shahin, M., Di Salle, A., & Márquez, G. (2021). Design, monitoring, and testing of microservices systems: The practitioners' perspective. Journal of Systems and Software, 182, 111061.
<https://doi.org/10.1016/j.jss.2021.111061>
- WORLD JUSTICE PROJECT. (2022). Rule of Law Index 2022. WJP.
<https://worldjusticeproject.org/rule-of-law-index/downloads/WJPIndex2022.pdf>
- Yáñez, A. (15 de junio de 2023). Minsait, clave en la digitalización de la justicia española. Diario de Sevilla. https://www.diariodesevilla.es/empresas-al-dia/Minsait-clave-digitalizacion-justicia-espanola_0_1802520461.html
- Zhang, B., & Soh, H. (2024). *Extract, Define, Canonicalize: An LLM-based Framework for Knowledge Graph Construction*. Conference on Empirical Methods in Natural Language Processing. arXiv. <https://doi.org/10.48550/arXiv.2404.03868>
- Zhang, L. (2020). *A Transfer Learning Approach for Handwritten Numeral Digit Recognition*. In Proceedings of the 3rd International Conference on Software Engineering and Information Management (ICSIM '20), New York, USA.
<https://doi.org/10.1145/3378936.3378970>
- Zhong, H., Xiua, C., Tu, C., Zhang, T., Liu, Z., & Sun, M. (2020). *How Does NLP Benefit Legal System: A Summary of Legal Artificial Intelligence*. En Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 5218–5230). Association for Computational Linguistics. 10.18653/v1/2020.acl-main.466

Anexos

Anexo 1: Delitos gestionados por la PSI - MININTER

Capítulo	Delito	Artículo
Contra la Fidelidad a la Funcion Militar Policial	Falsificacion o Adulteracion de Documentos Militar Policial	139
Contra la Seguridad Pública	Conducción en Estado de Ebriedad	274-A
Contra la Seguridad Pública	Fabricacion, Suministro o Tenencia de Materiales Peligrosos	279
Contra la Seguridad Pública	Tenencia Ilegal de Armas de Fuego y/o Municiones	279-G
Contra el Orden Migratorio	Trafico Illicito de Migrantes	303
Contra el Orden Migratorio	Tráfico Illicito de Migrantes Agravado	303-B
Contra la Administracion Pública	Violencia contra la Autoridad para Impedir el Ejercicio de sus Funciones	366
Contra la Administracion Pública	Resistencia o Desobediencia a la Autoridad	368
Contra la Administracion Pública	Violencia y Resistencia a la Autoridad	368
Contra la Administracion Pública	Omision, Rehusamiento o Demora de Actos Funcionales	377
Contra la Administracion Pública	Cohecho Generico	397
Contra la Administracion de Justicia	Falsa Declaracion en Procedimiento Administrativo	411
Contra la Fe Pública	Falsificacion y/o uso de Documento Falso	427
Contra la Fe Pública	Falsedad Idiologica	428
Contra la Fe Pública	Falsificacion de Sellos y Timbres Oficiales	434
Contra la Fe Pública	Falsedad Genérica	438
Contra la Fe Pública	Falsedad Genérica Agravada	438-A
Contra la Defensa Nacional	Colaboracion con Organizacion Ilegal	65

Nota. Elaboración Propia.

Anexo 2: Cuestionario para la población 1

N.	Pregunta	Alternativas				
		1	2	3	4	5
		Nunca	La mayoría de las veces no	Algunas veces si, algunas veces no	La mayoría de las veces si	Siempre
Dimensión 1: Recepción de expedientes judiciales						
1	Los expedientes judiciales que me asignan diariamente contienen la información completa.					
2	En el sistema SIE no he encontrado errores en la información cargada sobre los expedientes judiciales.					
3	No se me ha dificultado acceder a la información subida sobre los expedientes judiciales en el sistema SIE.					
4	El tiempo que toma cargar un expediente en el sistema SIE se realiza de manera rápida y toma muy poco tiempo.					
Dimensión 2: Elaboración y envío de recursos legales						
5	El tiempo promedio que me toma elaborar un recurso legal (formato de archivo PSI, apersonamiento, etc.) no demora más de 10 minutos.					
6	La carga de recursos legales, que tengo para elaborar durante el día, no me parece excesiva o pesada.					
7	La tasa de aceptación de los recursos legales que envío a otras entidades es buena.					

-
- 8 No he encontrado errores en la información y la redacción de los recursos legales posteriormente a su envío a otra entidad.
-

Dimensión 3: Planificación y participación en audiencias

- 9 Actualmente, las audiencias programadas me son comunicadas de manera oportuna.
-
- 10 No he tenido problemas para acceder a la información de las audiencias (fecha, hora, modalidad, lugar o enlace de videoconferencia).
-
- 11 He asistido a las audiencias que he tenido asignadas sin ningún tipo de inconveniente.
-
- 12 Durante mi participación en las audiencias no se me ha dificultado encontrar en el sistema la información del caso judicial para sustentar mi postura jurídica.
-

Dimensión 4: Revisión y seguimiento de expedientes

- 13 He podido realizar la identificación de tendencias o patrones en los expedientes sin problemas.
-
- 14 He podido realizar una revisión del estado de los expedientes de manera rápida y oportuna.
-
- 15 He podido informar a mi supervisor mi avance con respecto a la gestión de los expedientes judiciales que me han asignado.
-
- 16 He podido acceder a los puntos clave de cada expediente judicial para hacer el seguimiento respectivo.
-

Anexo 3: Fichas de registro de datos para la población 2

1. Datos del expediente

Código de expediente
Tipo de delito
Año de ingreso
Lugar

2. Registro de tiempos

2.1. Recepción de documentos

Tipo Documento	Nro documento	Inicio Recepción	Fin Recepción

2.2. Elaboración de recursos legales

Tipo Documento	Nro documento	Inicio Elaboración	Fin Elaboración

3. Cálculo de tiempos

3.1. Recepción de documentos

Tipo Documento	Nro documento	Tiempo Total de Recepción

3.3. Elaboración de recursos legales

Tipo Documento	Nro documento	Tiempo de Elaboración

4. Indicadores finales de tiempo

Indicador	Resultado
I1: Tiempo promedio de recepción de documentos	
I2: Tiempo promedio de elaboración de recursos legales	

Anexo 4: Validación de expertos

FICHA DE VALIDACIÓN DE INSTRUMENTO POR EXPERTO

TÍTULO DEL TRABAJO DE INVESTIGACIÓN / TESIS

Reingeniería del sistema web SIE para mejorar la gestión de expedientes judiciales en la Procuraduría Pública del Sector Interior – MININTER.

AUTORES:

Ortiz Abanto, Kevin Miguel

Pachas Lopez, Nabia Jasmin

INSTRUCCIONES: Coloque una "x" en el casillero correspondiente la valoración que experticia determine sobre las preguntas formuladas en el instrumento sobre la variable dependiente.

CRITERIOS	DESCRIPCIÓN	VALOR ASIGNADO POR EL EXPERTO									
		10	20	30	40	50	60	70	80	90	100
1. CLARIDAD	Está formado con el lenguaje adecuado										x
2. OBJETIVIDAD	Está expresado en conductas observables										x
3. ACTUALIDAD	Adecuado de acuerdo con el avance de la ciencia									x	
4. ORGANIZACIÓN	Existe una cohesión lógica entre sus elementos										x
5. SUFICIENCIA	Comprende los aspectos de la investigación										x
6. INTENCIONALIDAD	Adecuado para valorar los aspectos de la investigación										x
7. CONSISTENCIA	Basado en bases teóricas científicas										x
8. COHERENCIA	Hay correspondencia entre dimensiones indicadores e índices.										x
9. METODOLOGÍA	El diseño responde al propósito de la investigación.										x
10. PERTINENCIA	Es útil y adecuado para la investigación										x

PROMEDIO DE VALORIZACIÓN DE EXPERTO: 99

OBSERVACIONES POR EL EXPERTO:

Es un excelente instrumento de recolección de datos, se observa que se ha tenido el máximo cuidado en la elaboración de cada criterio

GRADO ACADÉMICO DEL EXPERTO: DOCTOR

INSTITUCIÓN DONDE LABORA: Universidad Nacional Mayor de San Marcos

APELLIDOS Y NOMBRES: Vega Huerta Hugo Froilán



DNI:06147737

Anexo 5: Instrumento 1- Matriz de Datos

Anexo 5.1- Matriz de datos pretest

UA	Edad	Género	Preg1	Preg2	Preg3	Preg4	Preg5	Preg6	Preg7	Preg8	Preg9	Preg10	Preg11	Preg12	Preg13	Preg14	Preg15	Preg16
1	36	F	1	2	2	1	1	1	3	2	1	1	1	2	1	1	4	2
2	42	M	2	3	3	2	2	1	4	3	3	3	3	2	2	3	3	3
3	35	F	5	5	5	3	3	3	5	5	5	5	5	5	5	5	5	4
4	50	M	3	4	4	3	3	3	4	4	4	3	4	3	3	4	4	4
5	33	F	2	2	3	2	2	2	3	3	2	2	2	3	2	3	3	2
6	57	M	3	1	4	3	3	2	4	4	4	3	4	3	3	3	4	3
7	40	M	1	2	2	2	2	1	3	2	1	2	2	1	1	2	2	2
8	53	M	2	2	3	2	2	2	4	3	3	2	3	2	3	5	3	3
9	39	M	1	1	2	1	2	1	3	2	1	2	1	1	1	2	2	2
10	49	F	2	3	2	2	2	2	3	4	2	2	2	2	4	2	3	3
11	42	M	5	2	3	2	2	2	5	3	3	3	4	3	3	3	4	3
12	46	M	4	3	5	2	3	3	5	4	4	4	5	4	5	4	5	5
13	47	M	3	3	4	2	3	2	4	3	3	3	4	3	3	3	4	3
14	34	M	3	3	4	3	2	3	4	4	4	3	3	4	4	3	4	4
15	45	F	3	3	4	3	3	3	3	4	4	3	3	3	3	3	3	3
16	48	M	5	4	5	3	3	2	5	5	5	5	5	5	3	5	5	5
17	55	F	4	4	4	3	4	3	5	4	4	4	5	4	4	4	5	5
18	47	M	4	4	4	3	3	2	4	3	4	3	4	4	4	4	4	4
19	41	M	3	3	4	3	3	2	4	3	4	3	3	4	4	4	4	4

20	55 F	4	3	4	4	4	3	5	4	4	4	5	4	4	4	5	5
21	45 M	5	5	5	3	4	4	4	5	5	5	5	5	5	3	5	5
22	37 M	3	3	3	3	3	3	3	4	3	2	4	4	3	3	4	3
23	36 M	3	3	4	3	3	2	4	3	4	3	4	4	4	3	4	4
24	36 F	3	3	3	2	3	3	4	5	4	3	4	4	3	3	4	4
25	58 M	1	1	2	1	2	1	3	2	1	2	1	1	1	2	2	2
26	51 F	5	4	5	4	4	3	4	5	3	3	4	5	3	5	5	3
27	58 F	3	5	3	2	2	2	4	4	5	3	4	3	3	4	4	5
28	51 F	3	3	3	2	3	3	4	4	4	3	4	4	4	3	4	4
29	51 M	4	4	5	4	4	4	4	4	4	4	5	5	5	5	5	4
30	40 F	4	4	4	3	4	3	5	4	4	4	5	4	4	4	5	5
31	52 F	3	3	4	4	3	3	5	4	4	2	4	4	4	4	5	5
32	33 M	3	3	4	4	4	3	4	4	4	4	4	3	4	4	4	5
33	31 M	5	5	5	3	3	3	5	5	5	5	5	5	5	5	5	4
34	26 F	5	5	5	3	4	3	5	5	5	5	4	4	5	5	5	5
35	28 M	1	2	2	1	1	1	3	3	2	2	2	2	3	1	3	2
36	43 M	5	5	5	4	3	3	5	5	5	5	5	5	5	5	4	4
37	42 F	4	5	4	4	4	3	5	5	5	5	4	5	4	4	5	5
38	60 F	5	5	4	4	3	3	5	5	5	4	5	5	5	4	5	4

Anexo 5.2- Matriz de datos posttest

UA	Edad	Género	Preg1	Preg2	Preg3	Preg4	Preg5	Preg6	Preg7	Preg8	Preg9	Preg10	Preg11	Preg12	Preg13	Preg14	Preg15	Preg16
1	36	F	1	2	2	3	3	2	3	2	2	2	3	2	2	3	3	2
2	42	M	3	2	4	3	4	3	4	3	4	2	4	3	3	4	4	3
3	35	F	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
4	50	M	5	5	5	5	5	3	5	4	4	4	4	5	4	5	4	4
5	33	F	3	3	3	4	3	4	4	3	3	3	3	3	2	3	3	3
6	57	M	3	3	4	4	3	3	4	4	4	3	4	3	3	4	4	4
7	40	M	2	1	2	3	3	2	2	2	2	2	3	2	2	3	3	3
8	53	M	3	4	4	4	4	2	4	3	4	4	4	4	3	4	3	4
9	39	M	3	3	2	2	3	2	3	3	3	3	2	2	3	2	2	2
10	49	F	3	3	3	4	4	2	4	3	4	3	4	4	3	3	3	3
11	42	M	4	4	4	5	4	5	4	4	4	4	4	5	4	4	4	5
12	46	M	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
13	47	M	4	4	4	5	4	5	4	4	4	5	4	3	4	4	5	5
14	34	M	5	2	3	5	5	4	5	5	3	5	5	5	3	5	5	5
15	45	F	3	3	3	4	4	4	4	3	4	3	3	4	2	3	4	4
16	48	M	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
17	55	F	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5	5
18	47	M	5	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5
19	41	M	5	5	5	5	5	5	5	4	5	4	5	4	4	5	5	4
20	55	F	5	5	5	5	5	4	5	4	5	5	5	4	5	5	5	5
21	45	M	5	5	5	5	5	4	5	4	5	4	5	4	4	4	4	4

22	37 M	3	3	4	4	3	3	4	4	4	3	4	3	3	4	4	4
23	36 M	5	5	5	5	5	5	5	4	5	5	5	5	5	5	5	5
24	36 F	5	5	5	5	5	3	5	5	4	4	4	4	5	5	4	4
25	58 M	3	2	2	2	2	3	2	2	2	2	3	3	3	3	3	3
26	51 F	5	5	5	5	5	4	5	4	5	5	5	4	5	5	5	5
27	58 F	4	4	4	5	4	5	4	4	4	5	4	4	3	4	5	5
28	51 F	5	5	5	5	5	5	5	4	5	5	5	5	5	5	5	5
29	51 M	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
30	40 F	5	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5
31	52 F	5	5	5	5	5	4	5	5	5	4	5	5	4	5	5	5
32	33 M	5	5	5	5	5	4	5	5	4	5	5	5	4	5	5	5
33	31 M	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
34	26 F	5	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5
35	28 M	3	3	2	2	3	2	2	2	3	3	2	3	3	2	2	3
36	43 M	5	5	5	5	5	3	5	4	4	4	4	5	4	5	4	4
37	42 F	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5	5
38	60 F	5	5	5	5	5	4	5	5	5	5	5	5	5	5	5	5

Anexo 6.2- Vista de variables

Archivo Editar Ver Datos Transformar Analizar Gráficos Utilidades Ampliaciones Ventana Ayuda Aplicación de búsqueda 6 días restantes en Prueba de ve											
	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida	Rol
1	Pretest_punt_D1	Númérico	4	2	Puntuación Promedio de Dimensión 1 (pretest)	Ninguna	Ninguna	12	Derecha	Escala	Entrada
2	Pretest_punt_D2	Númérico	4	2	Puntuación Promedio de Dimensión 2 (pretest)	Ninguna	Ninguna	12	Derecha	Escala	Entrada
3	Pretest_punt_D3	Númérico	4	2	Puntuación Promedio de Dimensión 3 (pretest)	Ninguna	Ninguna	12	Derecha	Escala	Entrada
4	Pretest_punt_D4	Númérico	4	2	Puntuación Promedio de Dimensión 4 (pretest)	Ninguna	Ninguna	12	Derecha	Escala	Entrada
5	Posttest_punt_D1	Númérico	4	2	Puntuación Promedio de Dimensión 1 (posttest)	Ninguna	Ninguna	12	Derecha	Escala	Entrada
6	Posttest_punt_D2	Númérico	4	2	Puntuación Promedio de Dimensión 2 (posttest)	Ninguna	Ninguna	12	Derecha	Escala	Entrada
7	Posttest_punt_D3	Númérico	4	2	Puntuación Promedio de Dimensión 3 (posttest)	Ninguna	Ninguna	12	Derecha	Escala	Entrada
8	Posttest_punt_D4	Númérico	4	2	Puntuación Promedio de Dimensión 4 (posttest)	Ninguna	Ninguna	12	Derecha	Escala	Entrada
9	Diff_D1	Númérico	4	2	Diferencia pretest y posttest para Dimensión 1	Ninguna	Ninguna	12	Derecha	Escala	Entrada
10	Diff_D2	Númérico	4	2	Diferencia pretest y posttest para Dimensión 2	Ninguna	Ninguna	12	Derecha	Escala	Entrada
11	Diff_D3	Númérico	5	2	Diferencia pretest y posttest para Dimensión 3	Ninguna	Ninguna	12	Derecha	Escala	Entrada
12	Diff_D4	Númérico	5	2	Diferencia pretest y posttest para Dimensión 4	Ninguna	Ninguna	12	Derecha	Escala	Entrada
13	Pretest_punt_gen	Númérico	8	2	Puntuación general (pretest)	Ninguna	Ninguna	10	Derecha	Escala	Entrada
14	Protest_punt_gen	Númérico	8	2	Puntuación general (posttest)	Ninguna	Ninguna	10	Derecha	Escala	Entrada
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											
25											
26											
27											
28											
29											
30											
31											

Visión general

Vista de datos

Vista de variables

Anexo 7: Instrumento 2 - Matriz de datos

Debido al gran volumen de información, la matriz de datos no se incluye íntegramente en el presente documento.

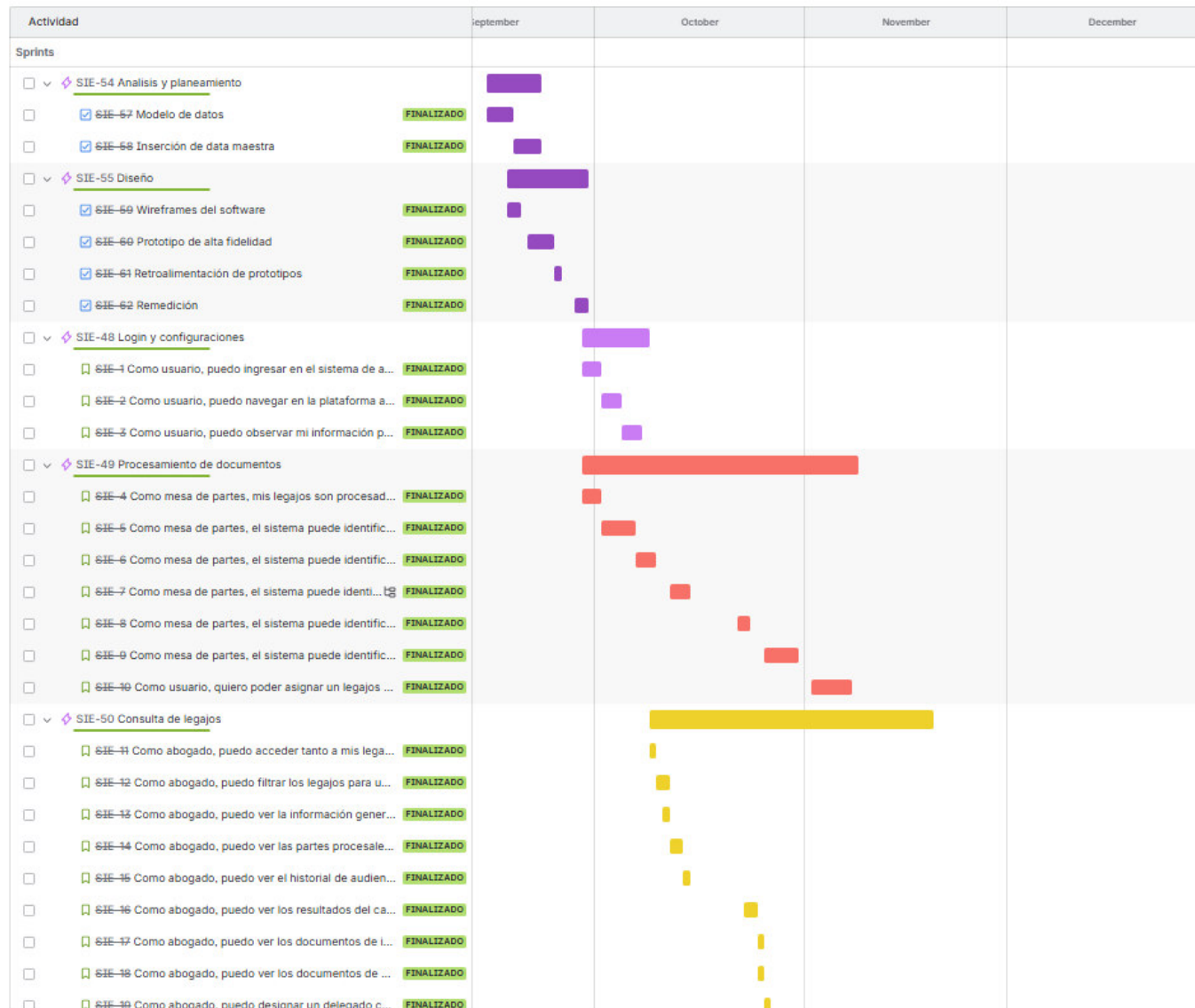
Puede ser consultada en el siguiente enlace:

https://docs.google.com/spreadsheets/d/1Jo0MAuh_ehHgyN_joEshqyclLoewodB--vtoGluTz7Y/edit?usp=sharing

Anexo 8: Presupuesto del Proyecto

Concepto	Mes 1			Mes 2			Mes 3			Mes 4			Mes 5		
	CU	PU	PT	CU	PU	PT	CU	PU	PT	CU	PU	PT	CU	PU	PT
Recursos Humanos			1.200,00			7.200,00			6.000,00			4.800,00			4.800,00
Diseñador	1	1.200,00	1.200,00	0	0,00	0,00	0	0,00	0,00	0	0,00	0,00	0	0,00	0,00
Desarrollador Frontend	0	0,00	0,00	1	2.400,00	2.400,00	1	2.400,00	2.400,00	1	2.400,00	2.400,00	1	2.400,00	2.400,00
Desarrollador Backend	0	0,00	0,00	1	2.400,00	2.400,00	1	2.400,00	2.400,00	1	2.400,00	2.400,00	1	2.400,00	2.400,00
Ingeniero de Inteligencia artificial	0	0,00	0,00	1	2.400,00	2.400,00	1	1.200,00	1.200,00	0	0,00	0,00	0	0,00	0,00
Servicios básicos			536,85			416,85			360,00			360,00			360,00
Luz	1	80,00	80,00	1	80,00	80,00	1	80,00	80,00	1	80,00	80,00	1	80,00	80,00
Internet	1	100,00	100,00	1	100,00	100,00	1	100,00	100,00	1	100,00	100,00	1	100,00	100,00
Agua	1	60,00	60,00	1	60,00	60,00	1	60,00	60,00	1	60,00	60,00	1	60,00	60,00
Asesoría de tesis	1	240,00	240,00	1	120,00	120,00	1	120,00	120,00	1	120,00	120,00	1	120,00	120,00
Cuenta Premium Figma	1	56,85	56,85	1	56,85	56,85	0	0,00	0,00	0	0,00	0,00	0	0,00	0,00
Servicios de Software			0,00			138,98			138,98			138,98			1.643,55
Servicio de base de datos en Azure	0	0,00	0,00	1	87,81	87,81	1	87,81	87,81	1	87,81	87,81	1	175,62	175,62
Servicios de hosting web	0	0,00	0,00	0	0,00	0,00	0	0,00	0,00	0	0,00	0,00	3	627,21	1.881,63
Servicio de OCR	0	0,00	0,00	1	51,17	51,17	1	51,17	51,17	1	51,17	51,17	1	153,51	153,51
Imprevistos	0	0,00	0,00	0	0,00	0,00	0	0,00	0,00	0	0,00	0,00	1	60,00	60,00
Subtotales			1.736,85			7.755,83			6.498,98			5.298,98			6.803,55
													TOTAL	S/.28.721,40	

Anexo 9: Cronograma del Proyecto



Actividad		November	December	January '25
Sprints				
☐ SIE-23 Como abogado, puedo descargar un document... FINALIZADO				
☐ SIE-24 Como abogado, puedo enviar por correo un do... FINALIZADO				
☐ SIE-25 Como abogado, puedo crear un documento de ... FINALIZADO				
☐ SIE-26 Como abogado, puedo editar la fase y el estad... FINALIZADO				
☐ SIE-27 Como abogado, puedo exportar el detalle de ... FINALIZADO				
☐ SIE-28 Como abogado, puedo exportar todos los legaj... FINALIZADO				
☐ SIE-29 Como abogado, puedo descargar un document... FINALIZADO				
☐ SIE-51 Recepción de legajos				
☐ SIE-30 Como mesa de partes, puedo una ventana que ... FINALIZADO				
☐ SIE-31 Como mesa de partes, puedo visualizar el cód... FINALIZADO				
☐ SIE-32 Como mesa de partes, puedo visualizar la infor... FINALIZADO				
☐ SIE-33 Como mesa de partes, puedo visualizar la infor... FINALIZADO				
☐ SIE-34 Como mesa de partes, puedo visualizar la infor... FINALIZADO				
☐ SIE-35 Como mesa de partes, puedo visualizar la infor... FINALIZADO				
☐ SIE-36 Como mesa de partes, puedo visualizar la infor... FINALIZADO				
☐ SIE-37 Como mesa de partes, puedo visualizar la infor... FINALIZADO				
☐ SIE-38 Como mesa de partes, puedo visualizar la infor... FINALIZADO				
☐ SIE-52 Reporte de legajos				
☐ SIE-39 Como procurador, puedo observar métricas so... FINALIZADO				
☐ SIE-40 Como procurador, puedo observar gráficos est... FINALIZADO				
☐ SIE-41 Como procurador, puedo visualizar el desempe... FINALIZADO				
☐ SIE-53 Audiencias				
☐ SIE-42 Como procurador, puedo observar las audienci... FINALIZADO				
☐ SIE-43 Como procurador, puedo editar las audiencias ... FINALIZADO				
☐ SIE-44 Como procurador, puedo insertar nuevas audie... FINALIZADO				
☐ SIE-45 Como abogado, puedo observar mis audiencia... FINALIZADO				
☐ SIE-46 Como abogado, puedo insertar nuevas audien... FINALIZADO				
☐ SIE-47 Como abogado, puedo editar mis audiencias FINALIZADO				
☐ SIE-56 Pruebas y despliegue				

Anexo 10: Prompts utilizados

```
meta_data:
  template: |
    Eres un extractor de entidades experto, tu objetivo es a partir de la descripción de un texto identificar las siguientes entidades:
    - clase_documento: debes elegir uno de los tipos encontrados en la lista proporcionada.
    - numero_documento: número del documento anexo a la clase de documento, este no siempre es entero, puede tener símbolos anexo.

    Las clases de documentos a identificar tanto para la clase como el número debe ser una de las siguientes:
    {clases}

    El texto del documento a identificar es el siguiente:
    [texto]: `{content}`.

    Tienes que tomar en consideración las siguientes indicaciones cuando extraigas la información:
    - No inventes algo que no se presente en el contenido del texto.
    - No inventes entidades ni valores no encontrados en el texto.
    - En caso consideres que no se evidencia la entidad en el texto el valor de esa entidad será -1.
```

```
dependencia_data:
  template: |
    Eres un extractor de entidades experto, tu objetivo es a partir de la descripción de un texto identificar la siguiente entidad:
    - dependencia: Índice de la dependencia de mininter identificada dentro del documento, puede ser únicamente una de las que es:

    Lista de Dependencias posibles:
    1. Policía Nacional del Perú. Tiene el acrónimo: PNP
    2. Intendencia Nacional de Bomberos del Perú. Tiene el acrónimo: INBP
    3. Superintendencia Nacional de Control de Servicios de Seguridad, Armas, Municiones y Explosivos de Uso Civil. Tiene el acrónimo: INSCASE
    4. Superintendencia Nacional de Migraciones. Tiene el acrónimo: Migraciones

    El texto del documento a identificar es el siguiente:
    [texto]:
    =====INICIO TEXTO DOCUMENTO=====
    `{content}`.
    =====FINAL TEXTO DOCUMENTO=====

    Tienes que tomar en consideración las siguientes indicaciones cuando extraigas la información:
    - No inventes algo que no se presente en el contenido del texto.
```