

**Plataforma de soporte para la evaluación de proyectos de investigación presentados  
ante el Comité de Ética de la Investigación del Hospital Universitario San Ignacio**



Facultad de ingeniería

Documento de Pruebas

Proyecto de grado

Jose Manuel Rodríguez  
Laura Valentina Ovalle  
Michael Joel González  
María Paula Cardona

Ingeniería de sistemas

Bogotá D.C.  
Octubre 2024

## HISTORIAL DE CAMBIOS

<b>Versión</b>	<b>Fecha</b>	<b>Sección del documento modificada</b>	<b>Descripción de cambios (corta)</b>	<b>Responsable (S)</b>
01	Mayo 2025	Todo el documento	Se creó la primer versión del documento	Jose Manuel Rodriguez Laura Valentina Ovalle Michael Joel Gonzalez Maria Paula Cardona

**Tabla 1: Historial de cambios**

## CONTENIDO

<b>HISTORIAL DE CAMBIOS.....</b>	<b>1</b>
<b>CONTENIDO.....</b>	<b>2</b>
<b>LISTA DE TABLAS.....</b>	<b>3</b>
<b>LISTA DE FIGURAS.....</b>	<b>3</b>
<b>1. INTRODUCCIÓN.....</b>	<b>4</b>
1.1 PROPÓSITO.....	4
1.2 OBJETIVOS.....	4
1.3 ALCANCE.....	5
<b>2. ESTRATEGIA DE PRUEBAS.....</b>	<b>5</b>
2.1 OBJETIVOS DE PRUEBAS.....	5
2.2 SUPUESTOS DE LAS PRUEBAS.....	5
2.3 NIVEL DE PRUEBAS.....	6
<b>3. ESTRATEGIAS DE EJECUCIÓN.....</b>	<b>6</b>
3.1 CRITERIOS DE ENTRADA Y SALIDA.....	6
3.1.1 Criterios de Entrada.....	6
3.1.2 Criterios de Salida.....	7
3.2 CICLO DE PRUEBAS.....	8
3.3 PROCESOS DE VALIDACIÓN.....	8
3.3.1 Métricas principales.....	8
<b>4. PRUEBAS REALIZADAS.....</b>	<b>9</b>
4.1 PRUEBAS DE INTEGRACIÓN.....	9
4.1.1 Módulo de Autenticación.....	10
4.1.2 Módulo de Caso.....	10
4.1.4 Módulo de Evaluación.....	11
4.1.5 Módulo de Inteligencia Artificial.....	12
4.1.6 Módulo de PDFs.....	12
4.1.7 Módulo de Prompts.....	13
4.1.8 Módulo de Correos.....	13
4.1.9 Módulo de Estadísticas.....	14
4.1.10 Módulo de Usuarios.....	14
4.2 PRUEBAS DE CARGA.....	15
4.2.1 Módulo de Casos.....	15
4.2.2 Módulo de Evaluación-LLM.....	16
4.2.3 Módulo Reset My Prompts.....	16
4.2.4 Módulo Prompts & IA-Config.....	17
4.2.5 Módulo Gestión de Usuarios.....	18
4.2.7 Módulo Autenticación (Login).....	20
4.2.8 Módulo de Estadísticas.....	21
4.3 PRUEBAS END TO END.....	22
4.3.1 Introducción a las Pruebas E2E.....	22
4.3.2 Configuración del Entorno de Pruebas.....	22
4.3.3 Configuración de Playwright.....	22

4.3.4 Cobertura de Pruebas.....	22
4.3.5 Flujo Evaluador.....	23
4.3.6 Flujo Investigador.....	25
4.4 PRUEBAS UAT.....	27
<b>5. EVIDENCIAS.....</b>	<b>32</b>
<b>6. REFERENCIAS.....</b>	<b>33</b>

## LISTA DE TABLAS

- Tabla 1: Historial de cambios
- Tabla 2: Tipos de pruebas
- Tabla 3: Criterios de entrada
- Tabla 4: Criterios de salida
- Tabla 5: Métricas principales de calidad ISO 25010
- Tabla 6: Resumen de métricas de rendimiento módulo casos
- Tabla 7: Resumen de métricas de rendimiento módulo evaluación LLM
- Tabla 8: Resumen de métricas de rendimiento módulo reset my prompts
- Tabla 9: Resumen de métricas de rendimiento módulo prompts y configuración...
- Tabla 10: Resumen de métricas de rendimiento módulo gestión de usuarios
- Tabla 11: Resumen de métricas de rendimiento módulo evaluaciones
- Tabla 12: Resumen de métricas de rendimiento módulo autenticación
- Tabla 13: Resumen de métricas de rendimiento módulo estadísticas
- Tabla 14: Resultados UAT

## LISTA DE FIGURAS

- Ilustración No.1 Cobertura del Módulo de Autenticación
- Ilustración No.2 Cobertura del Módulo de Caso
- Ilustración No.3 Cobertura del Módulo de Normas Éticas
- Ilustración No.4 Cobertura del Módulo de Evaluación
- Ilustración No.5 Cobertura del Módulo de Inteligencia Artificial
- Ilustración No.6 Cobertura del Módulo de Pdf
- Ilustración No.7 Cobertura del Módulo de Prompts
- Ilustración No.8 Cobertura del Módulo de Correos
- Ilustración No.9 Cobertura del Módulo de Estadísticas
- Ilustración No.10 Cobertura del Módulo de Usuarios
- Ilustración No.11 Vista general del escenario “cases-load...
- Ilustración No.12 Vista general del escenario “evaluationLLM-load...
- Ilustración No.13 Vista general del escenario “resetMyPrompts-...
- Ilustración No.14 Vista general del escenario “promptsAndConfigIA-load...
- Ilustración No.15 Vista del escenario “user-load”...
- Ilustración No.16 Vista del escenario “evaluation-load...
- Ilustración No.17 Vista del escenario “auth-load...
- Ilustración No.18 Vista del escenario “stats-load”...
- Ilustración No. 19 Distribución de respuestas para el módulo Autenticación.
- Ilustración No. 20 Distribución de respuestas para el módulo Subida de...
- Ilustración No. 21 Distribución de respuestas para el módulo Evaluación de...
- Ilustración No. 22 Distribución de respuestas para el módulo Configuración de...
- Ilustración No. 23 Distribución de respuestas para el módulo Proveedores y...
- Ilustración No. 24 Distribución de respuestas para el módulo Envío de...
- Ilustración No. 25 Distribución de respuestas para el módulo Creación de...
- Ilustración No. 26 Distribución de respuestas para el módulo Previsualización de...

# 1. INTRODUCCIÓN

El presente Plan de Pruebas describe con detalle las actividades de verificación y validación diseñadas para la *Plataforma de Asistencia del Comité de Ética del Hospital Universitario San Ignacio*. Su objetivo principal es confirmar que el sistema, ya implementado, satisface los requisitos funcionales y no funcionales definidos para agilizar la revisión ética de los consentimientos informados. Para lograrlo, se han definido y ejecutado cuatro frentes de prueba—Integración (Jest + Supertest), Carga/Estrés (Artillery), End-to-End (Playwright) y Pruebas de Aceptación de Usuario (UAT) con evaluadores externos—que abarcan desde la interacción de componentes individuales hasta la experiencia integral del usuario final. Cada fase cuenta con criterios formales de inicio y de aceptación, métricas de salida y procedimientos de reporte que garantizan un proceso transparente, reproducible y orientado a la mejora continua.

## 1.1 PROPÓSITO

El propósito fundamental de este Plan de Pruebas es establecer los procesos y las estrategias metodológicas necesarios para garantizar la máxima calidad y fiabilidad de la Plataforma de Asistencia para el Comité de Ética del Hospital Universitario San Ignacio.

Para ello, se define un marco integral de evaluación cuyo objetivo es demostrar que la plataforma satisfaga los requisitos funcionales definidos y mantenga un rendimiento aceptable con hasta 25 usuarios simultáneos —5 evaluadores y 20 investigadores— en un entorno de recursos limitados.

Asimismo, se describen los procesos sistemáticos y los criterios objetivos que guiarán cada fase del ciclo de pruebas, así como los procedimientos rigurosos para la validación de funcionalidades y desempeño.

## 1.2 OBJETIVOS

- Detectar y documentar fallos funcionales o de integración para permitir su corrección oportuna.
- Evaluar el desempeño y la escalabilidad midiendo tiempos de respuesta y consumo de recursos bajo una carga representativa de 25 sesiones simultáneas.
- Verificar los requisitos no funcionales—entre ellos usabilidad, seguridad, disponibilidad y accesibilidad—asegurando que el sistema se mantenga fiable en escenarios habituales y críticos.
- Recabar retroalimentación de usuarios reales; al menos diez participantes (familiares o colegas) deben calificar la experiencia como “fácil de usar”, validando así la aceptación y la satisfacción de la interfaz.

## 1.3 ALCANCE

El plan de pruebas comprende la revisión funcional y no funcional de ***al menos el 90 % de la plataforma***, cubriendo de manera explícita los itinerarios esenciales de sus dos actores: *Evaluador e Investigador*. Para ambos perfiles se validan la autenticación, la actualización de datos personales y el acceso a sus historiales. El rol Investigador incluye, además, el registro de nuevas cuentas, la creación, almacenamiento y descarga de casos con sus PDFs asociados; el rol Evaluador abarca la selección del modelo IA (Gemini o Groq), la configuración de prompts, la evaluación ética de documentos, la edición de normativas, el historial y la modificación de evaluaciones previas, la generación

automática de informes PDF y su envío por correo electrónico, así como la consulta de paneles estadísticos. Con esta cobertura se persigue el objetivo de demostrar que la solución cumple y satisface las expectativas de desempeño, seguridad y facilidad de uso del Hospital Universitario San Ignacio, garantizando un proceso de revisión ética ágil y fiable para todas las partes implicadas.

## 2. ESTRATEGIA DE PRUEBAS

Esta sección explica el enfoque elegido para verificar que la plataforma funciona correctamente, responde con buen rendimiento y brinda una experiencia satisfactoria a todos sus usuarios. Se describen los objetivos concretos, los supuestos bajo los cuales se ejecutan las pruebas y los niveles en los que se estructura la validación.

### 2.1 OBJETIVOS DE PRUEBAS

Verificar la adecuación funcional de los flujos críticos de Evaluador y Investigador:

- Autenticación y control de acceso;
- Configuración de proveedores de IA y límite de 15 peticiones/minuto;
- Operaciones CRUD de evaluaciones, normas y casos;
- Generación y manejo de informes PDF;
- Envío de correos y visualización de estadísticas.

y medir el rendimiento con cargas nominales de 25 usuarios y picos de hasta 50, evaluar los atributos de seguridad, compatibilidad, usabilidad (objetivo  $\geq 4$  en la escala Likert) y fiabilidad en contenedores, y registrar los hallazgos para su corrección antes de la entrega definitiva.

### 2.2 SUPUESTOS DE LAS PRUEBAS

Las pruebas se basan en ciertos supuestos que fijan el entorno, los recursos y los datos necesarios, asegurando que los escenarios reflejen condiciones de uso realistas. A continuación, se detallan los principales supuestos para la ejecución:

- El sistema se ejecuta en un entorno *Docker* local con frontend (puerto 3000), backend (puerto 8080), MongoDB y MinIO.
- Las *API keys* de Gemini y Groq y la conectividad SMTP están configuradas y disponibles.
- Se cuenta con los recursos de hardware y software (CPU, RAM y almacenamiento) suficientes para sostener pruebas de hasta 50 usuarios concurrentes.
- Los datos de prueba son sintéticos o proceden de documentos con permiso expreso de uso.
- El equipo de tesis (cuatro integrantes) realiza las pruebas de integración, carga y E2E; las pruebas de aceptación las ejecutan familiares y amigos bajo supervisión.

### 2.3 NIVEL DE PRUEBAS

Nivel	Herramienta / Técnica	Cobertura estimada	Atributos ISO 25010 clave
-------	-----------------------	--------------------	---------------------------

<b>Integración</b>	Jest + Supertest	90 % de endpoints	Adecuación funcional, seguridad
<b>Carga y Estrés</b>	Artillery	Flujo completo con 25–50 usuarios	Rendimiento, fiabilidad
<b>End-to-End</b>	Playwright (Chrome, Edge, Firefox, Safari, WebKit)	95 % de flujos críticos	Compatibilidad, usabilidad
<b>UAT</b>	Sesiones guiadas + Encuesta Likert	≥ 10 participantes externos	Usabilidad, satisfacción

**Tabla 2: Tipos de pruebas**

### 3. ESTRATEGIAS DE EJECUCIÓN

Esta sección describe las directrices que regirán la ejecución de las pruebas en la **Plataforma de Asistencia del Comité de Ética**. Primero se establecen los *criterios de entrada y de salida*, que delimitan las condiciones mínimas para iniciar y concluir la campaña de pruebas. Después se presenta el *ciclo de pruebas*, con las fases planificadas y la interacción entre ellas. Finalmente, se detallan los *procesos de validación* y las *métricas* que certificarán el cumplimiento de los requisitos funcionales y no funcionales.

#### 3.1 CRITERIOS DE ENTRADA Y SALIDA

##### 3.1.1 Criterios de Entrada

Característica	Especificación	Recursos asociados
Requisitos funcionales	Todo el <i>MVP</i> compila sin errores y queda accesible en Docker Compose.  Existe una cuenta Admin Evaluador pre-cargada.	Contenedores frontend:3000, backend:8080, MongoDB, MinIO.
Requisitos no funcionales	Límite gratuito de 15 peticiones/min a IA configurado; SMTP de prueba operativo.	Variables de entorno para Gemini, Groq y correo.
Datos de prueba	PDF de consentimiento autorizado cargado en la carpeta de pruebas.	Archivo PDF y colección Postman para creación masiva de usuarios Investigador.



Herramientas	Jest/Supertest, Artillery y Playwright instalados localmente en los portátiles del equipo (Node $\geq 20$ y Docker $\geq 24$ ).	Scripts npm test:int, test:load, test:e2e.
Seguridad	Revisión manual de terminales y logs confirma que no quedan tokens visibles tras el arranque.	Lista de comprobación impresa.

**Tabla 3: Criterios de entrada**

### 3.1.2 Criterios de Salida

Característica	Especificación	Resultados requeridos
Cobertura	$\geq 90$ % de los casos de prueba ejecutados y registrados.	Resumen combinado: <ul style="list-style-type: none"> <li>• Reporte de cobertura Jest (Imágenes)</li> <li>• Dashboards HTML y JSON de Artillery</li> <li>• Reporte HTML + videos de Playwright</li> </ul>
Defectos	Cero defectos <i>críticos</i> o <i>mayores</i> abiertos al cierre.	Resultados de las suites sin fallos y sin mensajes de error relevantes.
Rendimiento	Flujo completo soporta 25 usuarios nominales y picos de 50 sin errores 5xx.	Gráficas de latencia y throughput de Artillery.
Usabilidad	Media Likert $\geq 4/5$ con al menos 15 encuestas válidas.	Resumen estadístico (CSV).
Seguridad	Tasa de fugas de token = 0 %.	Bitácora de revisión manual sin hallazgos.

**Tabla 4: Criterios de salida**

### 3.2 CICLO DE PRUEBAS

- Planificación y preparación: Definir alcance de la ronda, asignar tareas y levantar contenedores.
- Ejecución de suites
  - Integración – Jest + Supertest sobre 90 % de endpoints.
  - Carga/Estrés – Artillery: 25 usuarios durante 5 min, luego rampa a 50 usuarios por 10 min.
  - End-to-End – Playwright *headless* en Chromium, Firefox y WebKit ejecutando los flujos críticos de Evaluador e Investigador.
- Registro de hallazgos: Resultados y evidencias se almacenan en un drive compartido.
- Re-test y regresión: Tras cada corrección se re-ejecutan los casos afectados.
- Cierre: Consolidar métricas, comparar contra los criterios de salida y emitir la decisión Aprobado/Re-trabajo.

### 3.3 PROCESOS DE VALIDACIÓN

Las suites se corren al menos dos veces: la primera para detección de fallos y la segunda para confirmar correcciones. La validación combina:

- Automática: Comparación de métricas generadas por las herramientas con los umbrales definidos
- Manual: Verificación visual de dashboards, confirmación de envío de correos y revisión de logs para asegurar ausencia de tokens.

#### 3.3.1 Métricas principales

Atributo ISO 25010	Indicador	Objetivo
Adecuación funcional	Cobertura funcional	$\geq 90 \%$
Rendimiento	Usuarios simultáneos sin error 5xx	25 nominal / 50 pico
Seguridad	Fugas de token	0 %
Usabilidad	Promedio Likert	$\geq 4$

Fiabilidad	Reinicios inesperados de contenedor	0
------------	-------------------------------------	---

**Tabla 5: Métricas principales de calidad ISO 25010**

## 4. PRUEBAS REALIZADAS

A lo largo del proyecto se ejecutan tres tipos de pruebas para demostrar que la plataforma cumple los requisitos de calidad y ofrece una experiencia fluida a los usuarios finales:

- Integración (Jest + Supertest). Verifican la interacción entre servicios REST y persistencia; cubren  $\approx 90\%$  de los endpoints.
- Carga / Estrés (Artillery). Evalúan el tiempo de respuesta y la estabilidad con cargas representativas de usuarios concurrentes.
- End-to-End (Playwright). Recorren los flujos críticos de Evaluador e Investigador en los cinco navegadores principales, validando usabilidad y compatibilidad.

Cada bloque de esta sección expone la metodología empleada, los escenarios configurados y las métricas recolectadas para certificar que el sistema alcanza los objetivos fijados en la Estrategia de Pruebas.

### 4.1 PRUEBAS DE INTEGRACIÓN

En esta sección se presentan los resultados de las pruebas de integración implementadas para validar la comunicación entre el servidor Express y la aplicación cliente, evaluando de manera exhaustiva los endpoints de las rutas API desarrolladas en TypeScript. Estas pruebas de integración fueron diseñadas para garantizar que los controladores del servidor respondan adecuadamente a las peticiones del cliente y procesen los datos conforme a los requisitos funcionales definidos.

Se empleó la librería Supertest junto con Jest para simular solicitudes HTTP sin requerir el despliegue del servidor en un entorno de producción. Esta aproximación permite evaluar cómo el servidor Express procesa y responde a las peticiones en un ambiente controlado, replicando los flujos de datos que ocurrirían en un entorno real. Al probar cada endpoint de forma aislada, se asegura que las operaciones fundamentales, como la creación, modificación, consulta y eliminación de recursos, operen correctamente y sin interferencias dentro del ecosistema de la aplicación.

### 4.1.1 Módulo de Autenticación

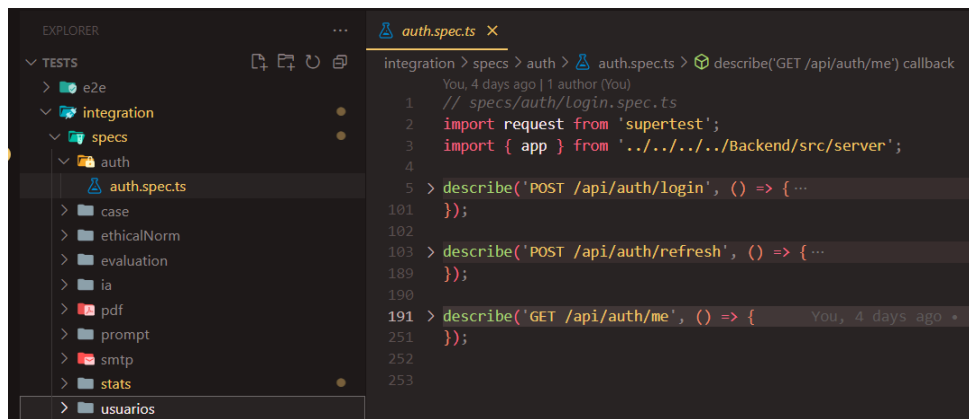


Ilustración No.1 Cobertura del Módulo de Autenticación

En la ilustración 1 se puede observar una cobertura del 100% en las pruebas de integración del módulo de autenticación, donde se validan todos los endpoints del “Auth.Controller.ts” asegurando el correcto funcionamiento de este.

### 4.1.2 Módulo de Caso

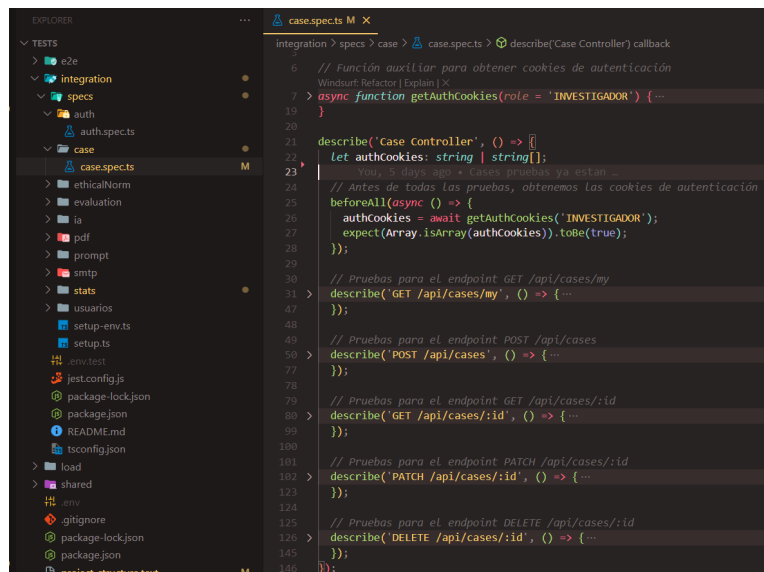
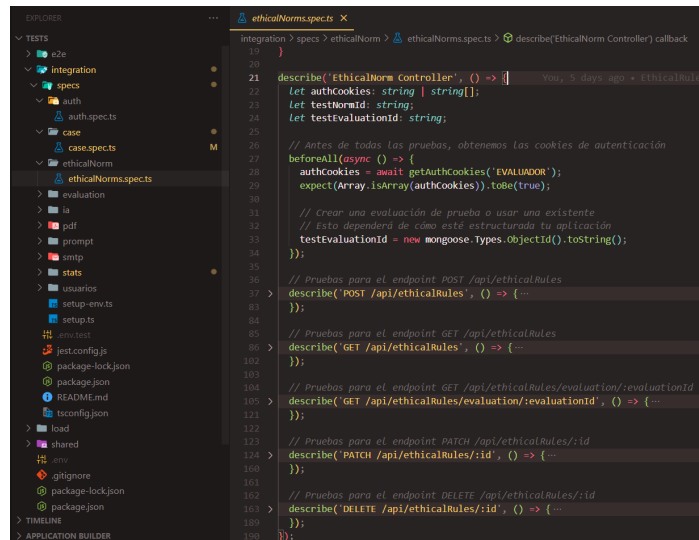


Ilustración No.2 Cobertura del Módulo de Caso

En la ilustración 2 se puede observar una cobertura del 100% en las pruebas de integración del módulo de autenticación donde se validan todos los endpoints de el “Case.Controller.ts” asegurando el correcto funcionamiento de este.

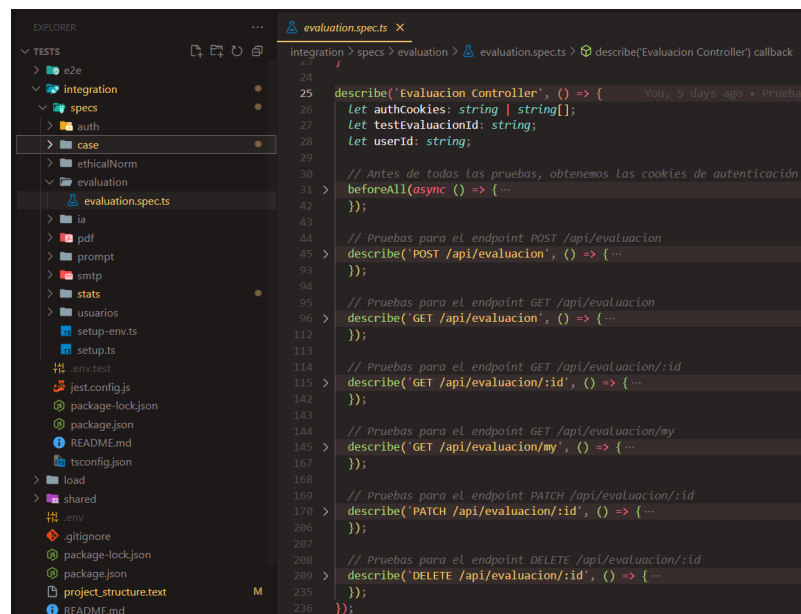
### 4.1.3 Módulo de Normas Éticas



**Ilustración No.3 Cobertura del Módulo de Normas Éticas**

En la ilustración 3 se puede observar una cobertura del 100% en las pruebas de integración del módulo de autenticación, donde se validan todos los endpoints del “EthicalRule.Controller.ts” asegurando el correcto funcionamiento de este.

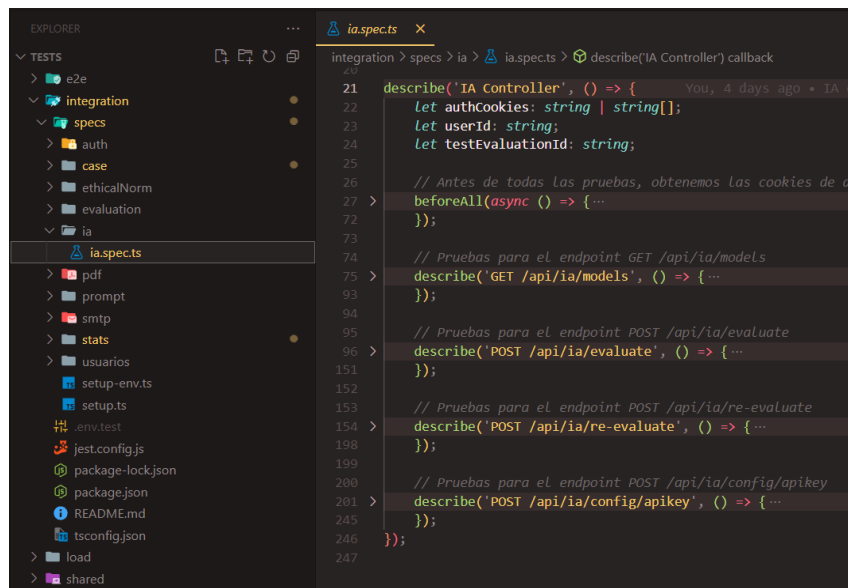
#### 4.1.4 Módulo de Evaluación



**Ilustración No.4 Cobertura del Módulo de Evaluación**

En la ilustración 4 se puede observar una cobertura del 100% en las pruebas de integración del módulo de autenticación, donde se validan todos los endpoints de él “Evaluation.Controller.ts” asegurando el correcto funcionamiento de este.

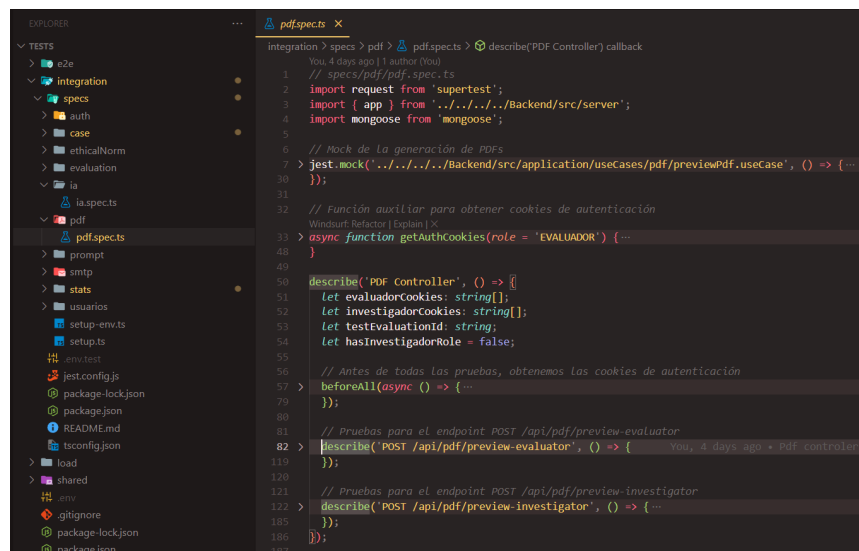
### 4.1.5 Módulo de Inteligencia Artificial



**Ilustración No.5 Cobertura del Módulo de Inteligencia Artificial**

En la ilustración 5 se puede observar una cobertura del 100% en las pruebas de integración del módulo de autenticación, donde se validan todos los endpoints de él “Ia.Controller.ts” asegurando el correcto funcionamiento de este.

### 4.1.6 Módulo de PDFs



**Ilustración No.6 Cobertura del Módulo de Pdf**

En la ilustración 6 se puede observar una cobertura del 100% en las pruebas de integración del módulo de autenticación donde se validan todos los endpoints de él “Pdf.Controller.ts” asegurando el correcto funcionamiento de este, para este caso se realizó un mockeo de la visualización del PDF para validar su funcionamiento.

### 4.1.7 Módulo de Prompts

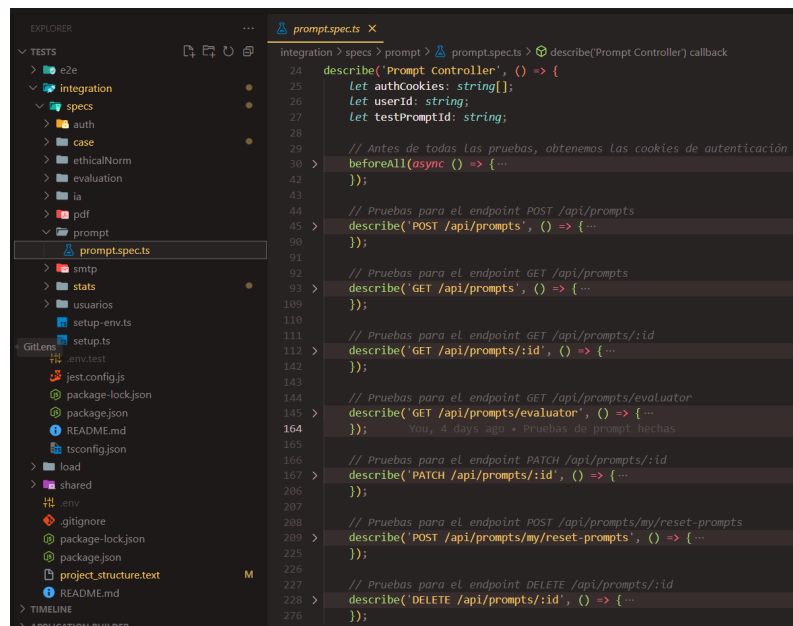


Ilustración No.7 Cobertura del Módulo de Prompts

En la ilustración 7 se puede observar una cobertura del 100% en las pruebas de integración del módulo de autenticación, donde se validan todos los endpoints de él “Prompts.Controller.ts” asegurando el correcto funcionamiento de este.

### 4.1.8 Módulo de Correos

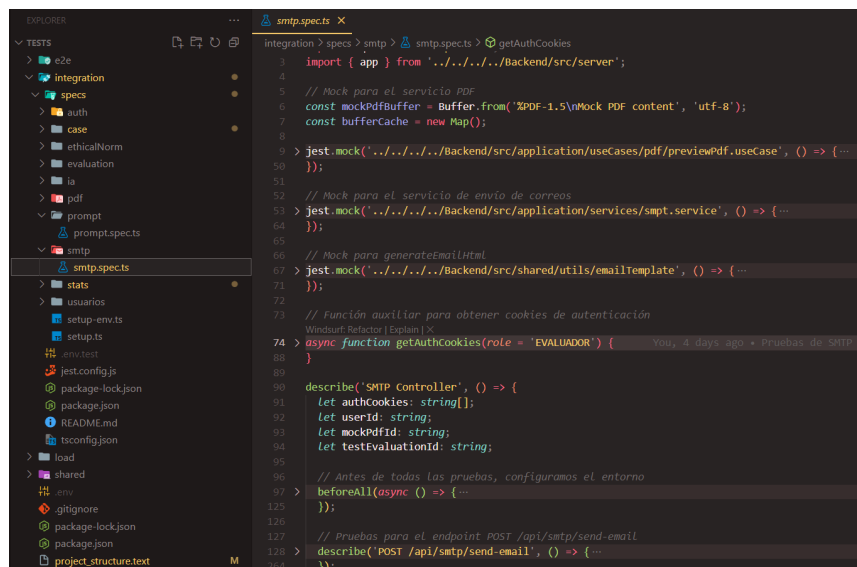


Ilustración No.8 Cobertura del Módulo de Correos

En la ilustración 8 se puede observar una cobertura del 100% en las pruebas de integración del módulo de autenticación donde se validan todos los endpoints de él “Smtplib.Controller.ts” asegurando el correcto funcionamiento de este, para este caso como el flujo de los correos depende de la funcionalidad del PDF esto fue mockeado.

### 4.1.9 Módulo de Estadísticas

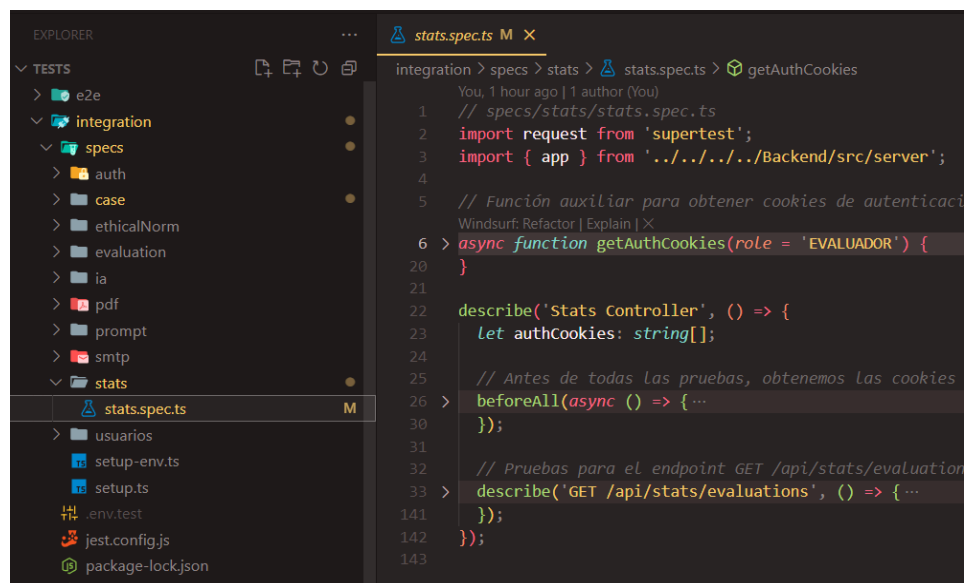


Ilustración No.9 Cobertura del Módulo de Estadísticas

En la ilustración 9 se puede observar una cobertura del 100% en las pruebas de integración del módulo de autenticación, donde se validan todos los endpoints de él “Stats.Controller.ts” asegurando el correcto funcionamiento de este.

### 4.1.10 Módulo de Usuarios

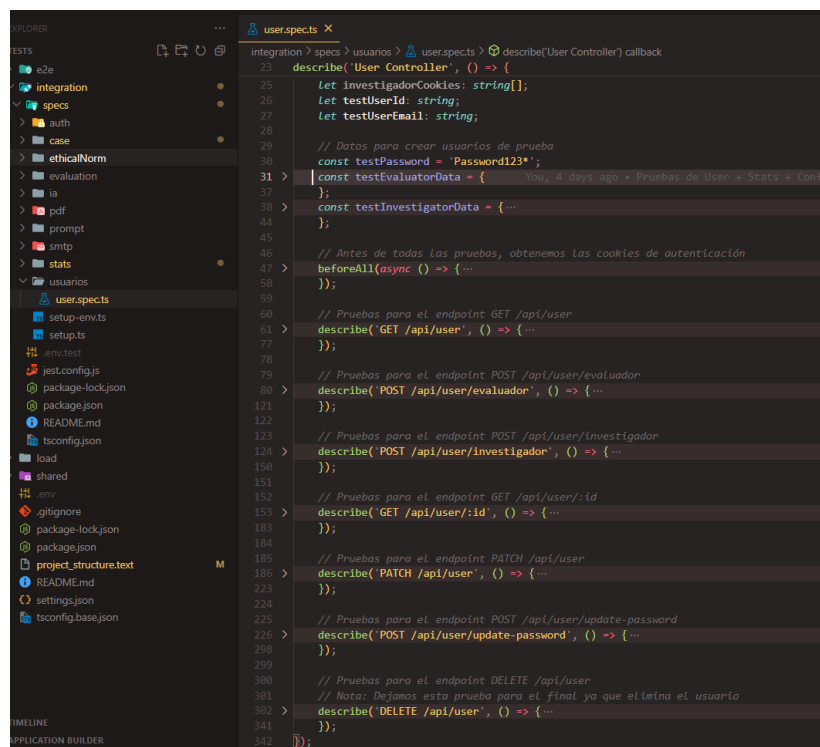


Ilustración No.10 Cobertura del Módulo de Usuarios



Por último, en la ilustración 10 se puede observar una cobertura del 100% en las pruebas de integración del módulo de autenticación, donde se validan todos los endpoints de él “Stats.Controller.ts” asegurando el correcto funcionamiento de este.

## 4.2 PRUEBAS DE CARGA

### 4.2.1 Módulo de Casos

Indicador global (50 VU, 250 peticiones)	Valor obtenido
Respuestas correctas (2xx)	100 %
Latencia media	211 ms
p50 / p90 / p95 / p99	86 ms / 714 ms / 805 ms / 1 200 ms
Throughput sostenido	≈ 2 req/s
Endpoint de mayor carga	Preview PDF → media ≈ 802 ms

Tabla 6: Resumen de métricas de rendimiento módulo casos

El módulo mantiene tiempos inferiores al segundo en el 95 % de las peticiones, incluso con 50 usuarios simultáneos. Las operaciones de negocio Create Case, Get My Cases, Login, Create Investigator responden de forma casi inmediata ( $p75 \leq 107$  ms).

Por otro lado, la generación y envío del PDF, que implica ensamblar el documento en el backend y remitirlo al frontend, se ejecuta con latencias propias de un procesamiento de archivos pero dentro del umbral previsto ( $< 2$  s p99), garantizando fluidez en la experiencia del Investigador.



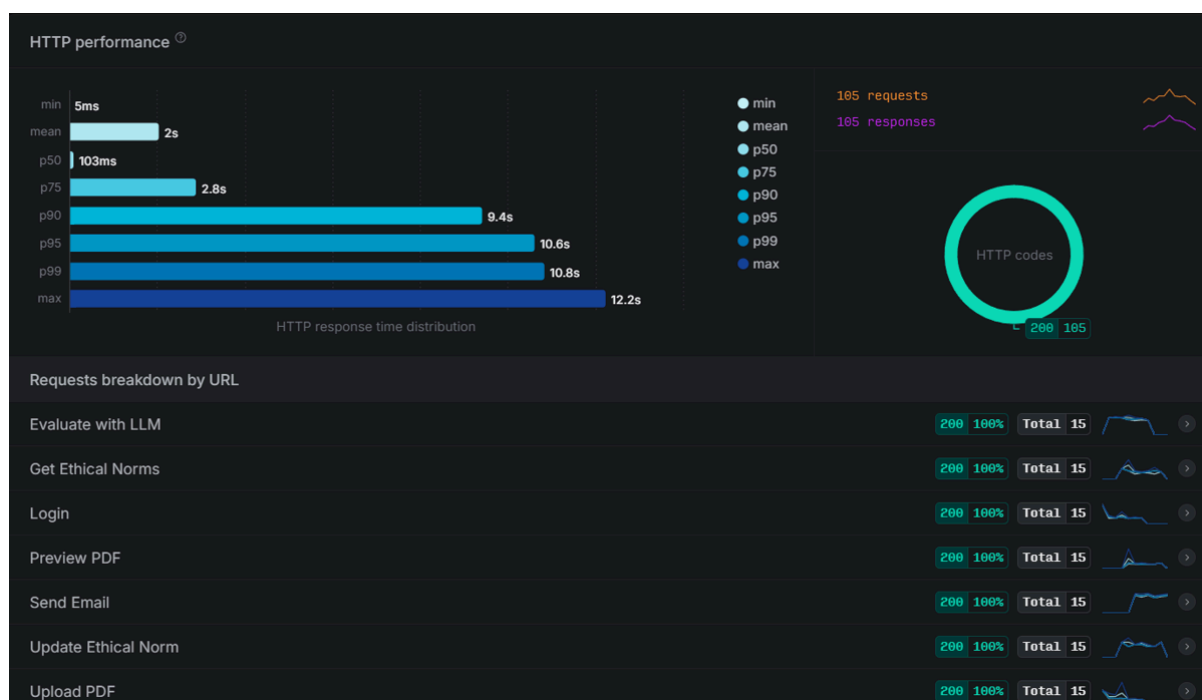
Ilustración No.11 Vista general del escenario “cases-load” en Artillery: curvas de VUs, latencia y distribución de códigos HTTP.

### 4.2.2 Módulo de Evaluación-LLM

Indicador global (15 VU, 105 peticiones)	Valor obtenido
Respuestas correctas (2xx)	100 %
Latencia media	2 041 ms
p50 / p90 / p95 / p99	102 ms / 9 417 ms / 10 618 ms / 10 832 ms
Throughput sostenido	≈ 2 req/s
Endpoint de mayor carga	Evaluate with LLM → media ≈ 10 311 ms

**Tabla 7: Resumen de métricas de rendimiento módulo evaluación LLM**

El flujo completo, autenticación, carga del PDF, consulta y actualización de normas éticas, evaluación con el LLM y envío de correo mantuvo una entrega correcta en el 100 % de los casos. Las operaciones intermedias (Login, Upload PDF, Get/Update Ethical Norm) responden casi en tiempo real ( $p75 \leq 106$  ms). Menos el endpoint de Evaluate with LLM, que delega cómputo al modelo de LLM externo, el cual presenta la mayor latencia debido al análisis que debe hacer del documento de consentimiento informado, sin embargo, su p99 ( $\approx 10,8$  s) continúa dentro del rango aceptable para tareas de inferencia pesada y no bloquea al resto de los usuarios concurrentes.



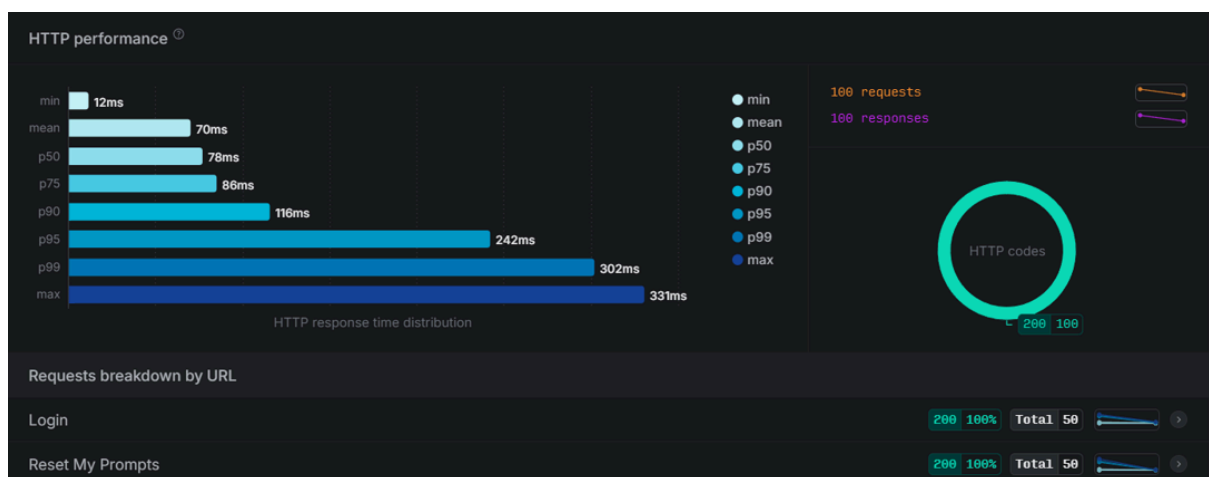
**Ilustración No.12 Vista general del escenario “evaluationLLM-load” en Artillery: curvas de VUs, latencia y códigos HTTP.**

### 4.2.3 Módulo Reset My Prompts

Indicador global (50 VU, 100 peticiones)	Valor obtenido
Respuestas correctas (2xx)	100 %
Latencia media	70 ms
p50 / p90 / p95 / p99	78 ms / 116 ms / 242 ms / 302 ms
Throughput sostenido	≈ 6 req/s
Endpoint de mayor carga	Login → media ≈ 102 ms

**Tabla 8: Resumen de métricas de rendimiento módulo reset my prompts**

El proceso de restablecimiento de prompts es muy ligero: tras la autenticación, la operación se limita a limpiar entradas en base de datos. Los valores muestran que el 99 % de las peticiones concluyen en < 0,3 s, garantizando una respuesta casi instantánea incluso cuando 50 usuarios solicitan el reseteo de forma simultánea. Esto valida la eficiencia del servicio de configuración personal ante picos de actividad.



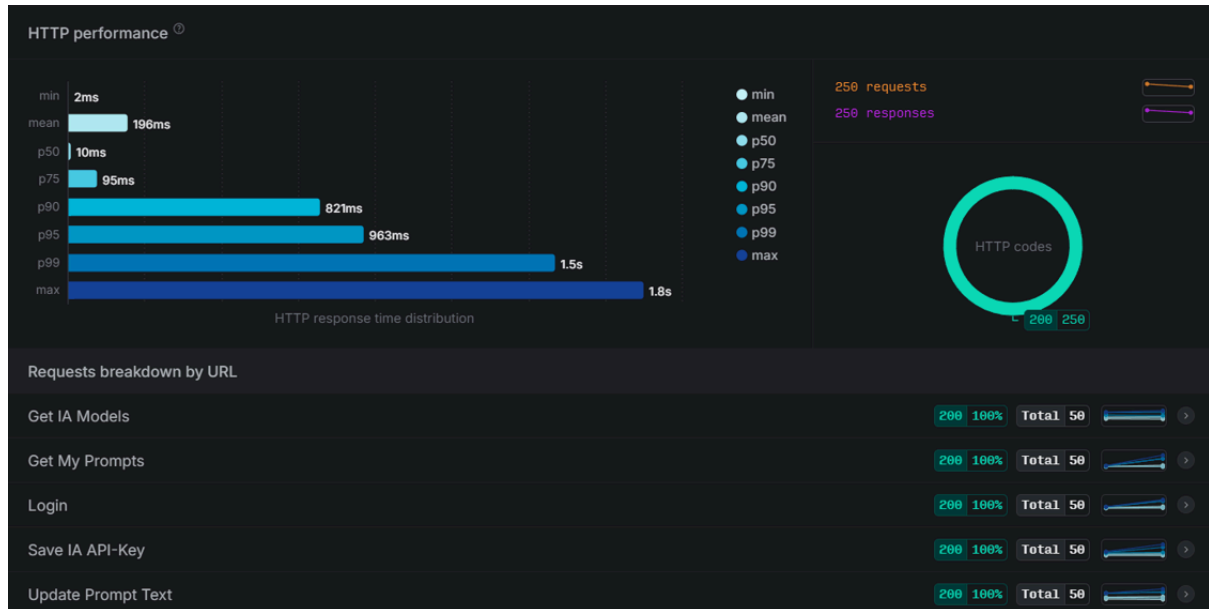
**Ilustración No.13 Vista general del escenario “resetMyPrompts-load” en Artillery.**

#### 4.2.4 Módulo Prompts & IA-Config

Indicador global (50 VU, 250 peticiones)	Valor obtenido
Respuestas correctas (2xx)	100 %
Latencia media	196 ms
p50 / p90 / p95 / p99	10 ms / 821 ms / 963 ms / 1 588 ms
Throughput sostenido	≈ 12 req/s
Endpoint de mayor carga	Get IA Models → media ≈ 861 ms

**Tabla 9: Resumen de métricas de rendimiento módulo prompts y configuración IA**

La mayoría de las operaciones (guardar API-Key, actualizar texto de prompt, obtener prompts) se resuelven en menos de 20 ms. El mayor costo está en Get IA Models, que implica contactar al proveedor externo y descargar el catálogo, pero aun así mantiene  $p95 < 1,3$  s. Bajo una presión de 50 usuarios concurrentes, la plataforma distribuye 12 peticiones por segundo sin degradar la experiencia: el tiempo de sesión promedio por usuario permanece alrededor del segundo.



**Ilustración No.14 Vista general del escenario “promptsAndConfigIA-load” en Artillery.**

#### 4.2.5 Módulo Gestión de Usuarios

Indicador global (50 VU, 300 peticiones)	Valor obtenido
Respuestas correctas (2xx / 201)	100 %
Latencia media	2 273 ms
p50 / p90 / p95 / p99	1 827 ms / 4 965 ms / 5 945 ms / 6 703 ms
Throughput sostenido	≈ 6 req/s
Endpoint de mayor carga	Update Password → media ≈ 4 787 ms

**Tabla 10: Resumen de métricas de rendimiento módulo gestion de usuarios**

Una sola prueba estresa todo el ciclo de vida de la cuenta: alta, primer login, edición de perfil, cambio de contraseña, segundo login y baja. Con 50 usuarios simultáneos no se registran fallos y la mitad de las operaciones concluye en menos de 2 s. La creación de usuarios, que implica validaciones y hashing inicial, se resuelve, en p75, antes de 3 s; el cambio de contraseña, tarea computacionalmente más intensa, marca el percentil 99 en ~6,7 s, sin embargo, esta tarea no bloquea el frontend, por lo que

continúa con una experiencia fluida y todavía dentro del rango aceptable para procesos criptográficos. Las acciones de lectura y eliminación (Update Profile, Delete User) permanecen por debajo de 1,5 s en p75, garantizando una experiencia ágil en la interfaz de administración.



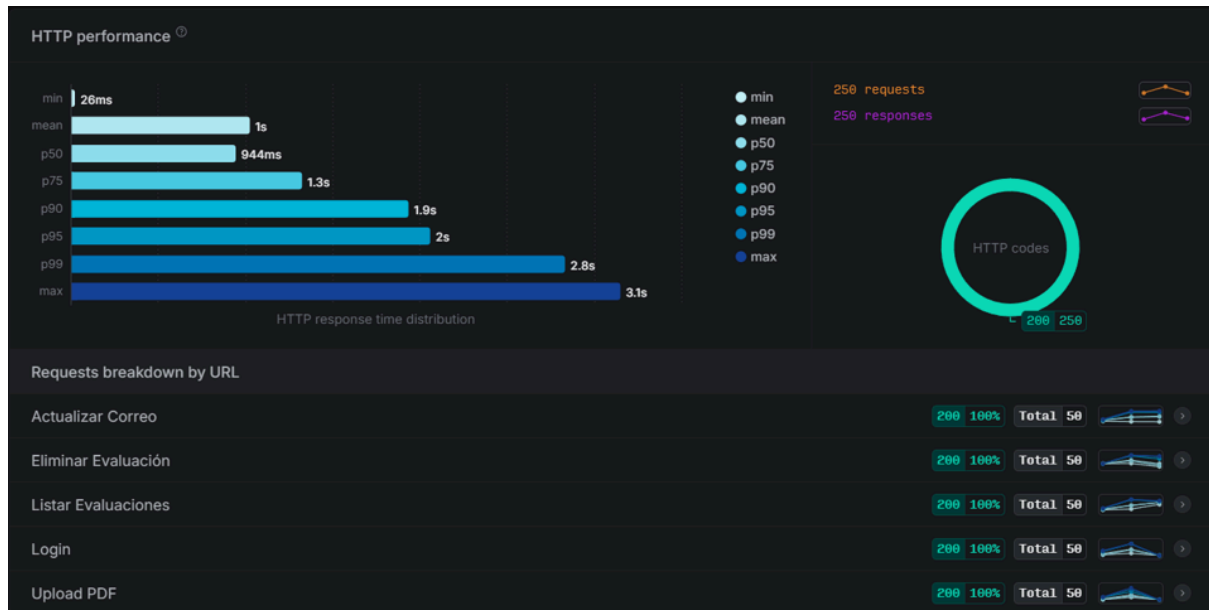
**Ilustración No.15** Vista del escenario “user-load” en Artillery: evolución de VUs, latencia y distribución de códigos HTTP.

4.2.6 Módulo Evaluaciones (CRUD sin LLM)

Indicador global (50 VU, 250 peticiones)	Valor obtenido
Respuestas correctas (2xx)	100 %
Latencia media	1 028 ms
p50 / p90 / p95 / p99	944 ms / 1 940 ms / 2 060 ms / 2 836 ms
Throughput sostenido	≈ 9 req/s
Endpoint de mayor carga	Eliminar Evaluación → media ≈ 1 274 ms

**Tabla 11: Resumen de métricas de rendimiento módulo evaluaciones**

El escenario reproduce la actividad cotidiana del evaluador: carga de documentos informados, historial de documentos, actualización de un documento y su eliminación. El servicio maneja 9 peticiones por segundo con el 95 % de las respuestas por debajo de 2,1 s. Las consultas (Listar Evaluaciones) y las actualizaciones simples (Actualizar Correo) se mantienen en el rango sub-segundo ( $p75 \leq 1,3$  s). La eliminación, que combina lógica de negocio y limpieza de ficheros, eleva la media a  $\sim 1,3$  s y fija el p99 en 2,8 s, sin comprometer la experiencia. Por lo que los códigos 2xx son constantes y la inexistencia de errores evidencian la estabilidad del módulo durante operaciones concurrentes.



**Ilustración No.16 Vista del escenario “evaluation-load” en Artillery.**

#### 4.2.7 Módulo Autenticación (Login)

Indicador global (50 VU, 50 peticiones)	Valor obtenido
Respuestas correctas (2xx)	100 %
Latencia media	86 ms
p50 / p90 / p95 / p99	81 ms / 99 ms / 118 ms / 125 ms
Throughput sostenido	≈ 3 req/s

Endpoint de mayor carga	Login → media ≈ 86 ms
-------------------------	-----------------------

**Tabla 12: Resumen de métricas de rendimiento módulo autenticación**

La prueba aísla el flujo de autenticación para medir únicamente el tiempo de verificación de credenciales y generación de la sesión JWT. Incluso con 50 accesos paralelos, el servidor responde en menos de 130 ms en el peor caso y resuelve la mitad de las solicitudes en ~80 ms, cifra cercana a la latencia de red local. Estos resultados confirman que la capa de seguridad hashing Bcrypt, firma y lectura del token, y no introduce cuellos de botella aun bajo concurrencia, proporcionando inicios de sesión prácticamente instantáneos.



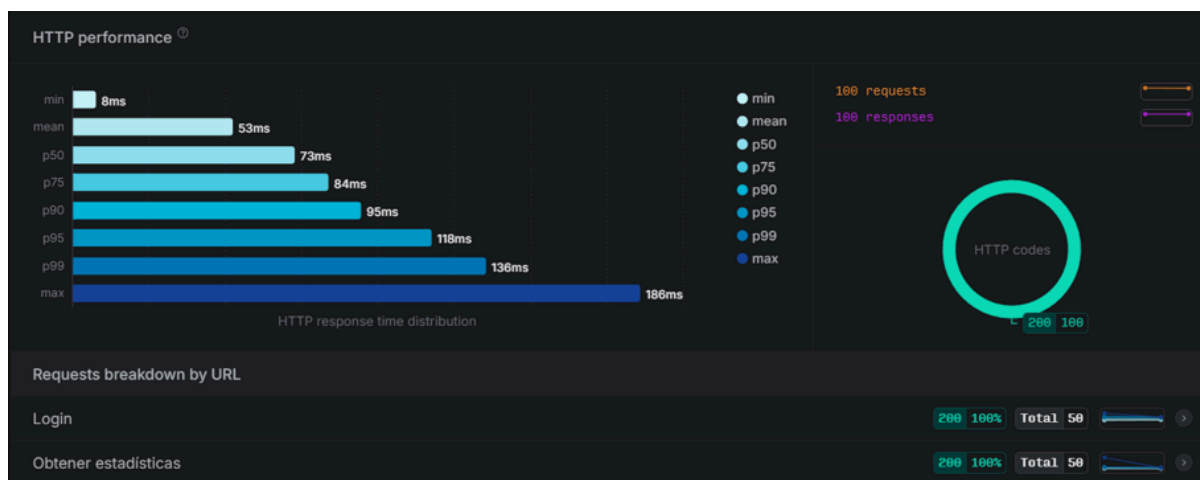
**Ilustración No.17 Vista del escenario “auth-load” en Artillery.**

#### 4.2.8 Módulo de Estadísticas

Indicador global (50 VU, 100 peticiones)	Valor obtenido
Respuestas correctas (2xx)	100 %
Latencia media	≈ 53 ms
p50 / p90 / p95 / p99	73 ms / 95 ms / 118 ms / 136 ms
Throughput sostenido	≈ 5 req/s
Endpoint de mayor carga	Login → media ≈ 90 ms

**Tabla 13: Resumen de métricas de rendimiento módulo estadísticas**

El flujo de consulta de estadísticas mantiene tiempos por debajo de 100 ms en el 90 % de las solicitudes, aun con cincuenta usuarios virtuales simultáneos. La operación crítica —Obtener estadísticas— responde casi instantáneamente (p75 ≈ 15 ms), lo que asegura que los cuadros y gráficas del dashboard se actualicen sin demoras perceptibles. El único punto que concentra algo más de latencia es el inicio de sesión (≈ 90 ms de media), valor que sigue estando dentro del rango recomendado para interacciones de autenticación.



**Ilustración No.18 Vista del escenario “stats-load” en Artillery.**

La constancia en los códigos 2xx y la ausencia total de errores corroboran la estabilidad del servicio. Lo que, en conjunto, los resultados confirman la solidez del backend bajo la carga objetivo.

## 4.3 PRUEBAS END TO END

### 4.3.1 Introducción a las Pruebas E2E

En esta sección se describe la implementación de pruebas de extremo a extremo (E2E) utilizando Playwright, una herramienta de automatización desarrollada por Microsoft que permite simular la interacción de usuarios reales con la aplicación web. Estas pruebas validan el funcionamiento completo del sistema desde la perspectiva del usuario final, verificando que todos los componentes trabajen correctamente en conjunto y que los flujos críticos de negocio operan según los requisitos establecidos.

### 4.3.2 Configuración del Entorno de Pruebas

Para la instalación y setup inicial se configuró Playwright en el proyecto utilizando TypeScript como lenguaje principal, aprovechando su compatibilidad nativa con este lenguaje. La instalación se realizó mediante npm, integrando las pruebas directamente en el repositorio del proyecto frontend para centralizar el código de testing.

### 4.3.3 Configuración de Playwright

El archivo playwright.config.ts se configuró para ejecutar las pruebas contra el servidor de desarrollo local, estableciendo la URL base y los comandos necesarios para levantar el entorno de testing. Se habilitaron artefactos esenciales como capturas de pantalla, videos y trazas para facilitar la depuración de fallos:

### 4.3.4 Cobertura de Pruebas

Las secuencias críticas validadas en las pruebas E2E se enfocaron en los flujos de usuarios que ofrecen mayor valor de negocio, incluyendo:

- Flujo Completo de Investigador
- Flujo Completo del Evaluador



#### 4.3.5 Flujo Evaluador

Esta prueba de extremo a extremo válida el flujo completo de un usuario evaluador en el sistema, desde la autenticación hasta la gestión de documentos y configuraciones. A continuación se detalla cada fase del proceso:

##### *Fase 1: Autenticación*

- La prueba inicia con el proceso de login del usuario evaluador:
- Navegación a la página de autenticación (/auth)
- Ingreso de credenciales (correo: admin-husi@gmail.com, contraseña: Password123\*)
- Validación del acceso exitoso al sistema

##### *Fase 2: Configuración de Inteligencia Artificial*

- Una vez autenticado, el usuario accede a las configuraciones de IA:
- Navegación al módulo de configuración de IA
- Selección del proveedor de IA (gemini)
- Cambio del modelo de IA a gemini-2.0-flash
- Confirmación de los cambios realizados
- Reinicio de prompts del sistema mediante confirmación con palabra clave REINICIAR

##### *Fase 3: Personalización de Apariencia*

- El flujo continúa con la configuración de preferencias visuales:
- Acceso al módulo de apariencia
- Modificación de configuraciones de tema/diseño
- Actualización de preferencias de visualización

##### *Fase 4: Gestión de Cuenta de Usuario*

- Se valida la funcionalidad de actualización de perfil:
- Modificación del nombre de usuario con timestamp único (Admin-`{Date.now()}`)
- Confirmación de cambios en la cuenta
- Validación de persistencia de datos

##### *Fase 5: Gestión de Archivos*

- Esta fase prueba el sistema de carga y administración de documentos:
- Carga de archivo PDF (IFC 489 INICIAL.pdf) mediante drag & drop
- Gestión de archivos duplicados
- Validación del proceso de subida exitosa

##### *Fase 6: Administración y Filtrado de Documentos*

- Se evalúan las funcionalidades de gestión de documentos:
- Configuración de vistas personalizadas (mostrar/ocultar columnas)
- Aplicación de filtros por estado (Pendiente, En curso, Evaluado, Aprobado, No aprobado)
- Limpieza de filtros aplicados
- Funcionalidad de búsqueda en tiempo real

##### *Fase 7: Edición y Evaluación de Documentos*

- Esta sección valida el flujo crítico de evaluación:

- Edición de metadatos del documento (correo del estudiante, tipo de error)
- Proceso de evaluación mediante selección de criterios específicos
- Validación de múltiples criterios de evaluación del documento

#### *Fase 8: Comunicación y Notificaciones*

- Se prueba el sistema de notificaciones:
- Envío de resultados de evaluación por correo electrónico
- Selección de categoría de feedback (Ortografía)
- Inclusión de mensajes personalizados
- Confirmación de envío exitoso

#### *Fase 9: Gestión del Historial*

- Validación de funcionalidades de seguimiento:
- Acceso al historial de evaluaciones
- Funcionalidad de reevaluación de documentos
- Navegación entre módulos del sistema

#### *Fase 10: Eliminación de Datos*

- Se valida el proceso de eliminación segura:
- Eliminación de documentos con confirmación mediante palabra clave ELIMINAR
- Validación de la persistencia de cambios

#### *Fase 11: Cierre de Sesión*

- La prueba finaliza validando el proceso de logout:
- Acceso al menú de usuario
- Cierre seguro de sesión
- Validación de redirección apropiada

### **Características Técnicas de la Prueba**

- *Elementos Validados*
  - Autenticación y autorización de usuarios administrativos
  - Gestión completa de configuraciones del sistema
  - Operaciones CRUD sobre documentos y metadatos
  - Flujos de comunicación y notificaciones
  - Integridad de datos en operaciones críticas
- *Patrones de Testing Implementados*
  - Uso de localizadores semánticos (getByRole, getByText) para mayor estabilidad
  - Generación de datos únicos con timestamps para evitar conflictos
  - Confirmaciones explícitas para operaciones destructivas
  - Validación de estados antes de proceder con acciones

Esta prueba E2E proporciona una cobertura integral del sistema, validando que todos los componentes funcionen correctamente en conjunto y que los flujos críticos de negocio operen según los requisitos establecidos.

### 4.3.6 Flujo Investigador

Esta prueba de extremo a extremo válida el flujo completo de un nuevo usuario investigador desde el registro hasta la gestión integral de casos de investigación. A continuación se detalla cada fase del proceso:

#### *Fase 1: Registro de Nuevo Usuario*

- La prueba inicia con el proceso de registro de un usuario nuevo:
- Navegación a la página de autenticación (/auth)
- Cambio a la pestaña de "Registro"
- Ingreso de datos personales (nombre: Juan, apellido: Rodriguez)
- Generación de correo único con timestamp (testuser\_`\${Date.now()}`@example.com)
- Establecimiento de contraseña (Password1\*)
- Confirmación del registro exitoso

#### *Fase 2: Creación de Nuevo Caso de Investigación*

- Una vez registrado, el usuario crea un caso:
- Acceso al módulo de creación de casos
- Inicio del proceso de configuración del documento

#### *Fase 3: Configuración de Cabecera del Documento*

- Se valida la funcionalidad de configuración inicial:
- Establecimiento de versión del documento (1)
- Asignación de código identificador (123)
- Selección de fecha específica (Monday, May 19th)

#### *Fase 4: Sección de Introducción*

- Configuración de información básica del proyecto:
- Definición del nombre del proyecto (Trabajo)
- Especificación de instituciones participantes (Trabajo)

#### *Fase 5: Información General del Estudio*

- Esta fase valida la configuración de parámetros críticos:
- Selección de nivel de riesgo (Riesgo mínimo)
- Definición de población objetivo (Adultos mayores de 18 años)
- Clasificación del tipo de estudio (Intervención (psicosocial))
- Completado de 14 campos de información detallada, todos con el valor Trabajo

#### *Fase 6: Información de Patrocinadores e Investigadores*

- Se valida la gestión de datos de contacto:
- Datos del patrocinador: nombre (Sara), compañía (Apple), dirección (Apple)
- Información del investigador principal: nombre (Juan), celular (321054826), género (Masculino)
- Datos del director de investigaciones: nombre (Laura), contacto (juan@gmail.com)

#### *Fase 7: Sección de Autorización*

- Configuración de información de autorización:
- Nombre del estudio (Gripe)

- Datos del investigador principal con género (Masculino)
- Información de contacto completa (celular: 321054826, teléfono: 321054826, extensión: 57)
- Nombre del presidente del comité (Juan)

*Fase 8: Pre Visualización y Manipulación del PDF*

- Se valida la funcionalidad del visor de documentos:
- Generación de previsualización del PDF
- Navegación entre páginas del documento
- Funciones de manipulación: rotación múltiple (4 rotaciones de 90 grados)
- Control de zoom: cambio a 50% y regreso a 100%
- Validación de todas las herramientas de visualización

*Fase 9: Guardado y Gestión del Caso*

- Validación de persistencia y administración:
- Guardado exitoso del caso creado
- Acceso al menú de opciones del caso
- Visualización del PDF generado
- Cierre correcto del visor

*Fase 10: Personalización de Vistas*

- Se prueba la funcionalidad de configuración de interfaz:
- Configuración de columnas visibles (Versión, Nombre Proyecto)
- Validación de cambios en tiempo real
- Reversión de configuraciones

*Fase 11: Eliminación Segura de Casos*

- Validación del proceso de eliminación:
- Acceso al menú de eliminación
- Confirmación mediante palabra clave ELIMINAR
- Validación de eliminación exitosa

*Fase 12: Actualización de Perfil de Usuario*

- La prueba finaliza con la gestión de cuenta:
- Acceso a configuraciones de cuenta
- Modificación de datos personales (nombre: Jose, apellido: Prueba)
- Confirmación de cambios realizados

***Características Técnicas de la Prueba***

- Elementos Validados
  - Registro completo de usuarios con validación de datos únicos
  - Creación integral de casos de investigación con múltiples secciones
  - Gestión de formularios complejos con validaciones específicas
  - Funcionalidades de pre visualización y manipulación de documentos
  - Operaciones CRUD completas sobre casos y perfiles de usuario
- Patrones de Testing Implementados

- Generación de datos únicos con timestamps para evitar conflictos
  - Navegación secuencial a través de formularios multi-paso
  - Validación de persistencia de datos en múltiples secciones
  - Confirmaciones explícitas para operaciones críticas
  - Uso de localizadores semánticos para mayor estabilidad
- Flujos de Negocio Críticos
    - Onboarding completo de nuevos usuarios
    - Proceso de creación de documentos de investigación
    - Gestión de metadatos y configuraciones personalizadas
    - Validación de integridad de documentos generados

Esta prueba E2E proporciona una cobertura completa del flujo de usuario desde el registro hasta la gestión avanzada de casos, asegurando que todos los componentes del sistema funcionen correctamente en un escenario de uso real y completo.

#### 4.4 PRUEBAS UAT

En esta sección describimos las pruebas de aceptación realizadas con usuarios externos (compañeros de universidad y amigos) que usaron la aplicación por primera vez. El objetivo fue comprobar la usabilidad y claridad de cada módulo bajo condiciones reales de uso.

##### Metodología

Invitamos a 21 participantes sin experiencia técnica previa en el desarrollo de la plataforma. Tras una breve introducción y demo, les pedimos que interactúan libremente con cada uno de los siguientes módulos:

- Autenticación
- Subida de archivos
- Evaluación de archivos
- Configuración de prompts
- Selección de proveedores y modelos de IA
- Envío de resultados
- Creación de casos
- Previsualización de PDF

A continuación respondieron un cuestionario de 15 ítems en escala Likert de 1 (Muy en desacuerdo) a 5 (Muy de acuerdo) valorando facilidad de uso y claridad de la interfaz de cada módulo.

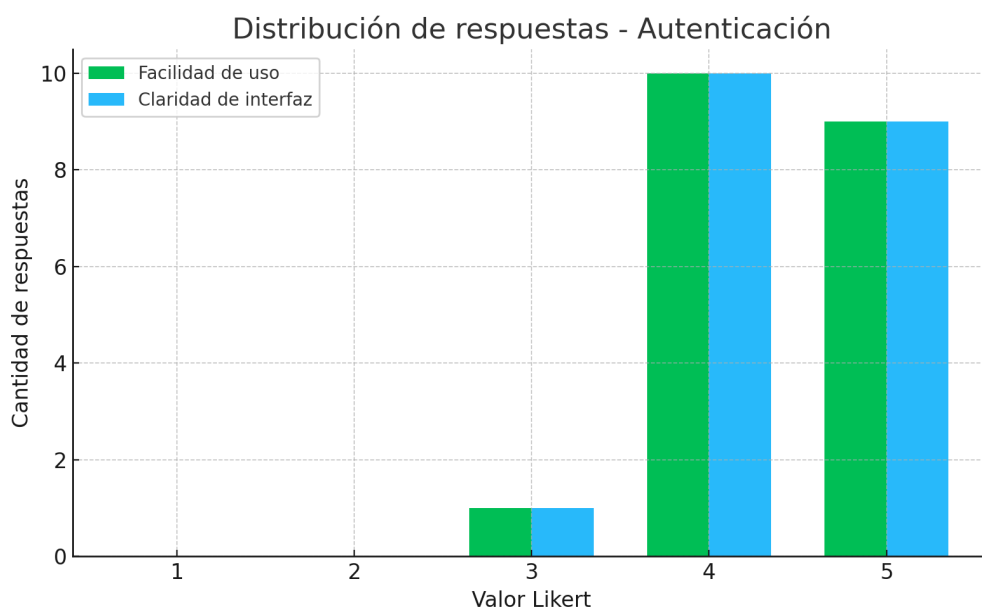
##### Resultados

A continuación se muestra un resumen de las puntuaciones medias y desviaciones estándar obtenidas por ítem:

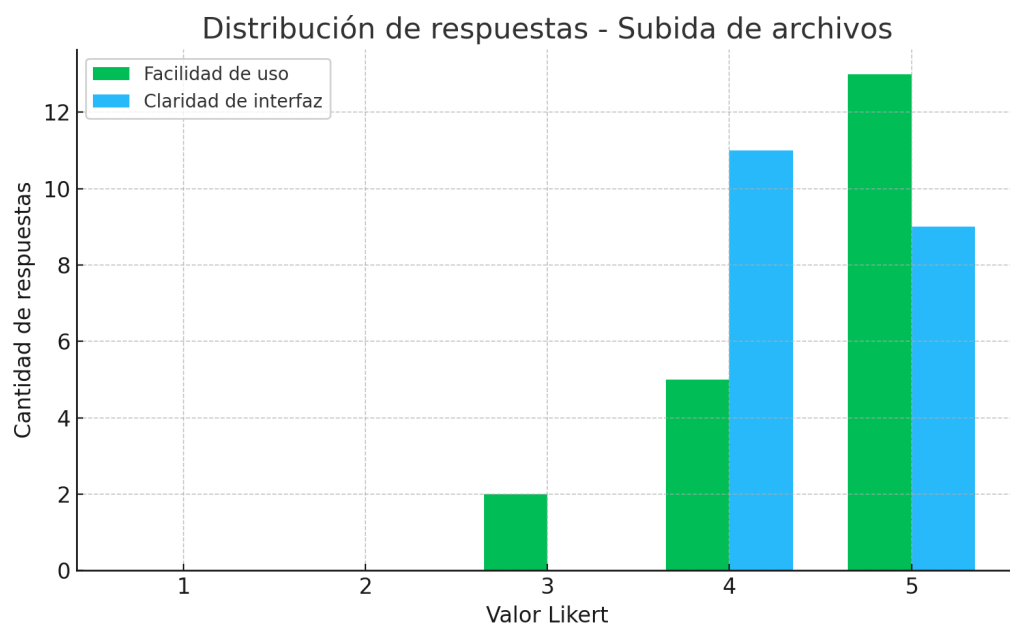
Módulo	Ítem	Media	$\sigma$ (DE)
Autenticación	Facilidad de uso	4,4	0,58

	Claridad de interfaz	4,4	0,58
<b>Subida de archivos</b>	Facilidad de uso	4,55	0,67
	Claridad de interfaz	4,45	0,5
<b>Evaluación de archivos</b>	Facilidad de uso	4,25	0,77
	Claridad de interfaz	4,3	0,78
<b>Configuración de prompts</b>	Facilidad de uso	4,2	0,51
	Claridad de interfaz	4,35	0,65
<b>Selección de proveedores y modelos</b>	Facilidad de uso	4,5	0,67
	Claridad de interfaz	4,65	0,57
<b>Envío de resultados</b>	Facilidad de uso	4,55	0,59
	Claridad de interfaz	4,5	0,59
<b>Creación de casos</b>	Facilidad de uso	4,65	0,48
	Claridad de interfaz	4,45	0,67
<b>Previsualización de PDF</b>	Facilidad y claridad (ítem único)	4,5	0,5

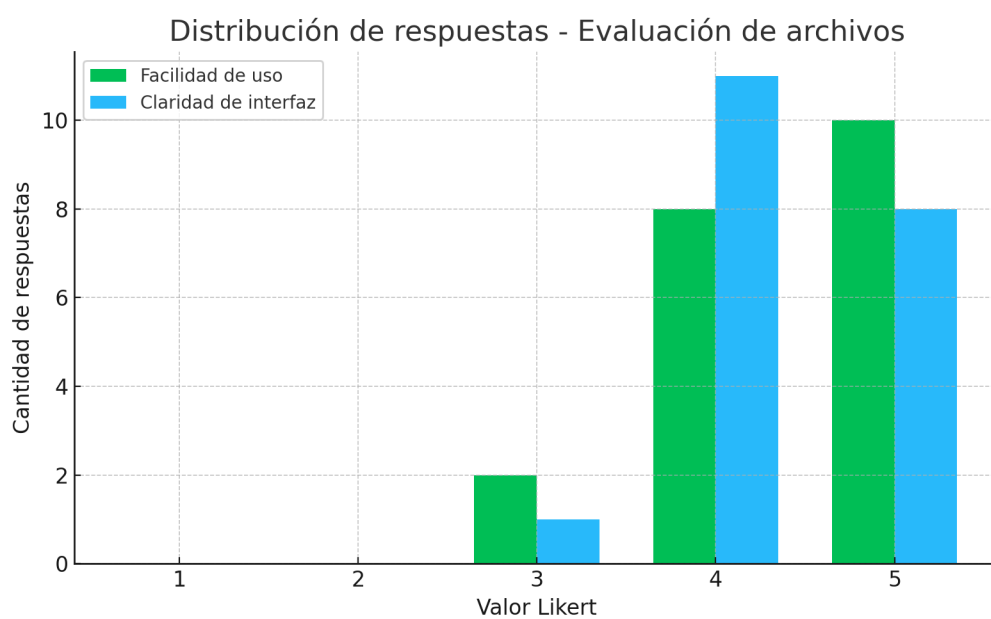
**Tabla 14: Resultados UAT**



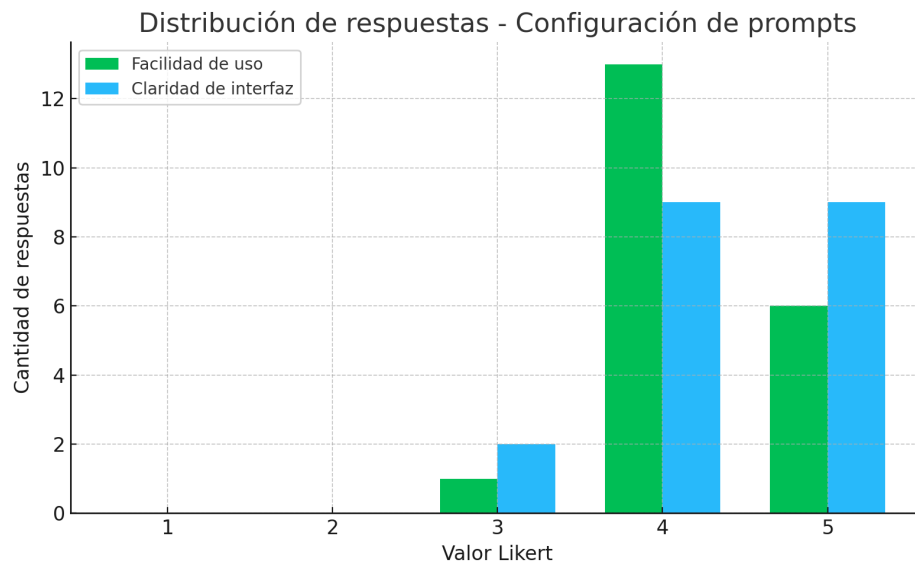
**Ilustración No. 19 Distribución de respuestas para el módulo Autenticación.**



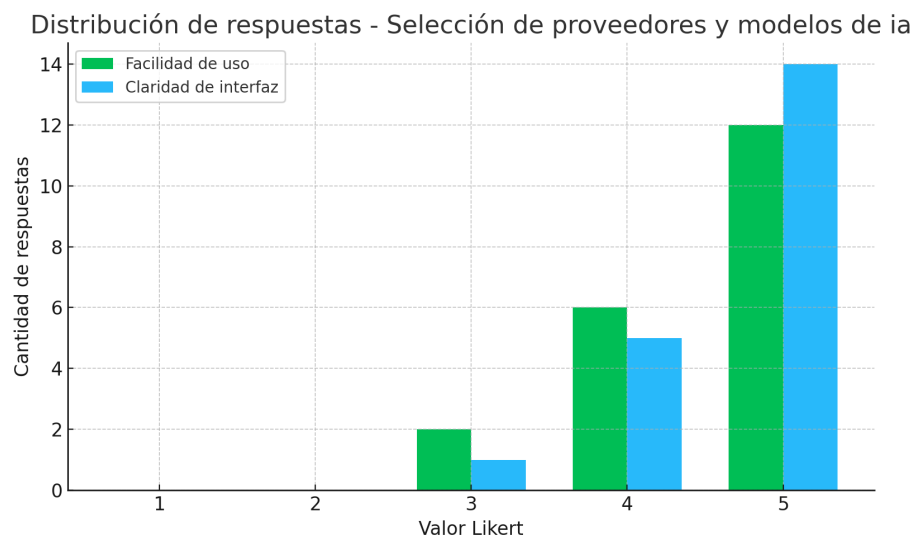
**Ilustración No. 20 Distribución de respuestas para el módulo Subida de archivos.**



**Ilustración No. 21 Distribución de respuestas para el módulo Evaluación de archivos.**

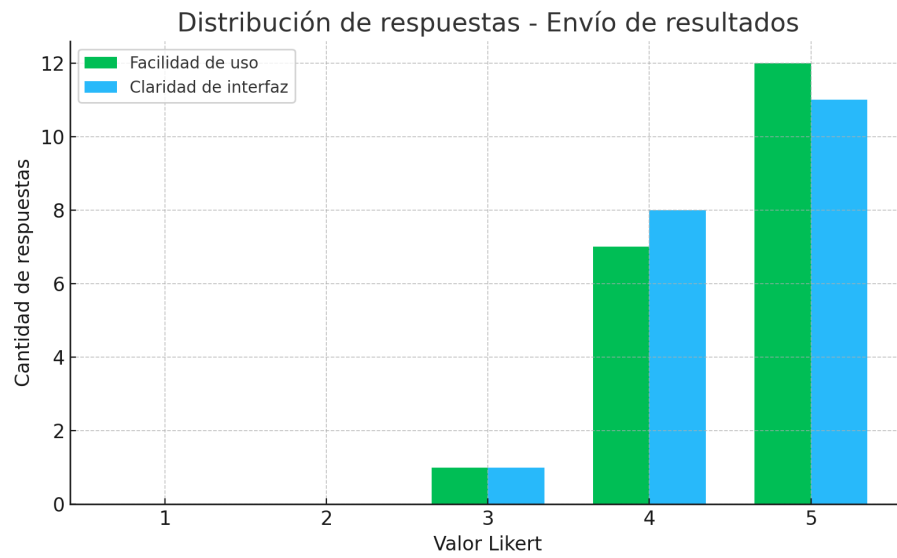


**Ilustración No. 22 Distribución de respuestas para el módulo Configuración de prompts.**

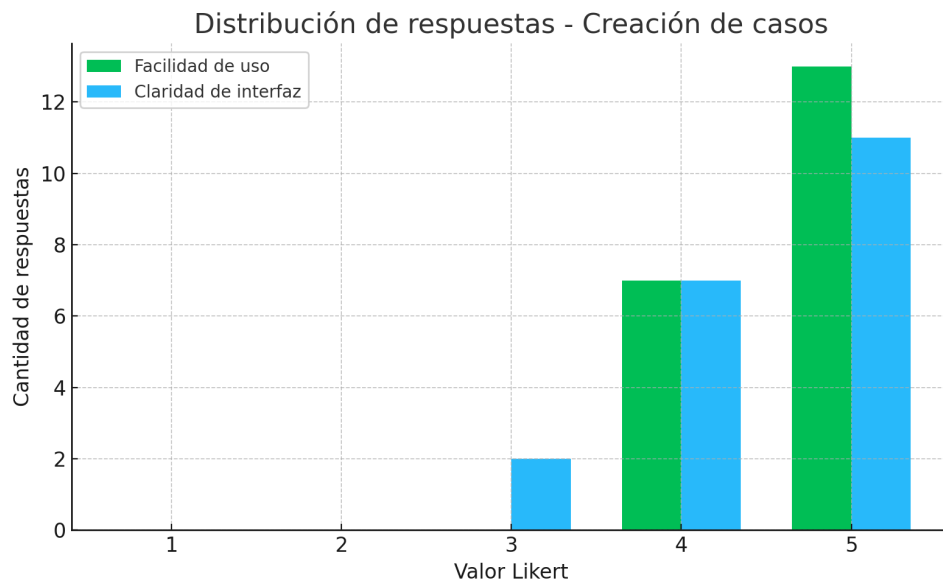


**Ilustración No. 23 Distribución de respuestas para el módulo Proveedores y modelos de ia.**

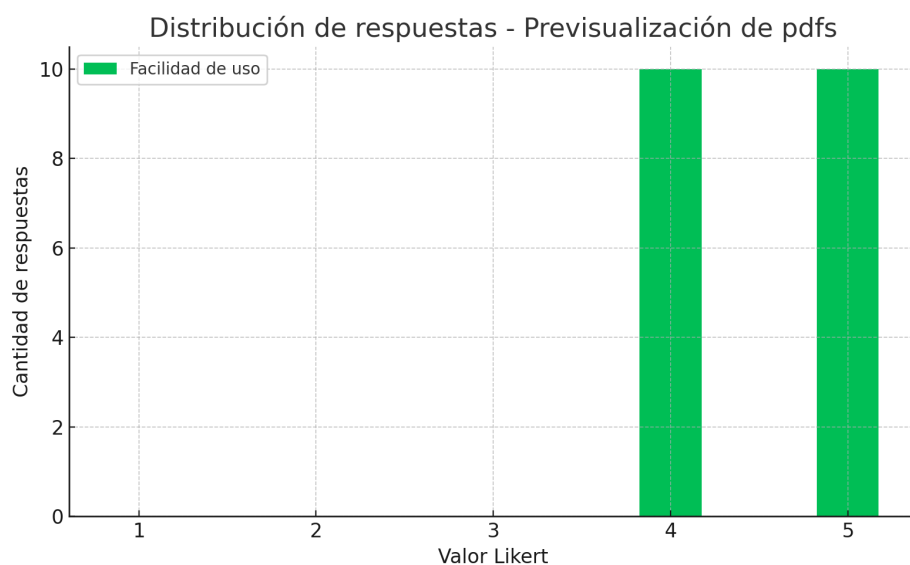




**Ilustración No. 24** Distribución de respuestas para el módulo Envío de resultados.



**Ilustración No. 25** Distribución de respuestas para el módulo Creación de casos.



**Ilustración No. 26 Distribución de respuestas para el módulo Previsualización de pdfs.**

En síntesis, la prueba de aceptación con usuarios dejó muy buenas sensaciones. Todas las valoraciones se mantuvieron por encima de 4,2 lo que, en práctica, significa que la gran mayoría de personas se sintió “de acuerdo” o “muy de acuerdo” con la facilidad y la claridad de cada módulo. Una constancia que sugiere que no hay puntos débiles evidentes ni curvas de aprendizaje pronunciadas.

## 5. EVIDENCIAS

Para todas las evidencias más detallada se podrá revisar los anexos, en esta carpeta se encontraran, carpetas con todas las fotos del proceso automatizado de Playwright, un archivo Excel con todos los resultados de las pruebas UAT, todas las estadísticas e información detallada de las pruebas de carga hechas a través de Artillery y por último las fotos de los resultados de las pruebas de Integración realizadas con Jest y SuperTest.

## 6. REFERENCIAS


[1] Jest. "Documentation." *Jest · Delightful JavaScript Testing*. Accessed: May 25, 2025. [Online]. Available: <https://jestjs.io/docs/>

[2] Ladj's. "SuperTest Repository." *GitHub*. Accessed: May 25, 2025. [Online]. Available: <https://github.com/ladjs/supertest>

[3] Artillery Inc. "Documentation." *Artillery: Modern load testing for developers*. Accessed: May 25, 2025. [Online]. Available: <https://www.artillery.io/docs/>

[4] Microsoft. "Documentation." *Playwright - Fast and reliable end-to-end testing for modern web apps*. Accessed: May 25, 2025. [Online]. Available: <https://playwright.dev/docs/>

## 7. ANEXOS

1. Carpeta de Evidencia de Pruebas:  Evidencia Pruebas