

PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR

FACULTAD DE INGENIERÍA

CARRERA DE INGENIERÍA EN SISTEMAS DE INFORMACIÓN



TRABAJO DE TITULACIÓN

INTEGRACIÓN DE HERRAMIENTAS DE INTELIGENCIA ARTIFICIAL
GENERATIVA EN LA GESTIÓN DE PROYECTOS: UN ESTUDIO APLICADO

ZARGARAN FAREDI ARMAN

QUITO, JUNIO DE 2025

AGRADECIMIENTOS

Esta tesis es un reflejo del apoyo incondicional y del amor desbordante que recibí de mi familia y amistades durante este desafiante recorrido académico. Estoy profundamente agradecido a mis padres, quienes cultivaron mi curiosidad y respaldaron mis esfuerzos educativos desde el principio. Quiero expresar mi gratitud a mis profesores de la Pontificia Universidad Católica del Ecuador, cuya pasión por la enseñanza y compromiso con la excelencia me inspiraron en cada paso. Sus enseñanzas en el aula y sus consejos fuera de ella han sido fundamentales para mi crecimiento académico y personal. Un agradecimiento muy especial a mi directora de tesis, Inge. Suyana Arcos, cuya orientación experta, comentarios constructivos y estímulo constante guiaron esta investigación hasta su finalización. Su mentoría ha sido pilar esencial de este trabajo. También agradezco a mis amigos, quienes brindaron distracciones cuando las necesitaba y ánimos cuando parecía imposible continuar. Las conversaciones y el apoyo inquebrantable de mis compañeros de universidad enriquecieron mi experiencia de investigación, haciendo que este camino no solo fuera posible, sino también muy gratificante.

RESUMEN

La presente investigación aborda la integración de herramientas de Inteligencia Artificial Generativa (GenAI) en la gestión de proyectos mediante un estudio aplicado en un caso real de recolección y etiquetado de imágenes. Ante la carencia de evidencia práctica sobre el impacto de estas tecnologías en entornos operativos, se propuso un marco de evaluación comparativo que incluyó cinco soluciones de GenAI aplicadas de manera secuencial durante cinco semanas con un equipo de nueve miembros. El objetivo principal fue determinar cuál de estas herramientas ofrecía el mayor valor en términos de eficiencia, calidad y productividad, utilizando indicadores clave de desempeño relativos a supervisión (tiempos de planificación y revisión) y a la ejecución operativa (volumen de imágenes procesadas, velocidad y tasa de error). La metodología combinó el análisis cuantitativo de KPIs con encuestas de percepción al supervisor y al equipo, validando las recomendaciones generadas por cada herramienta bajo condiciones de uso idénticas. Los resultados mostraron variaciones significativas en productividad (hasta +17 %), reducción de errores (hasta -51 %) y ahorro de tiempo de supervisión (hasta -47 %). De manera integradora, ChatGPT Pro (GPT-4o) emergió como la opción más equilibrada, liderando en múltiples métricas y recibiendo las mejores valoraciones de usuario. Finalmente, se propone un modelo híbrido IA-gestor de proyectos y buenas prácticas para maximizar el aprovechamiento de GenAI en la gestión de proyectos reales.

ABSTRACT

This study examines the integration of Generative Artificial Intelligence (GenAI) tools into project management through an applied case involving a weekly image collection and labeling workflow. Despite growing theoretical interest, there is limited empirical evidence on GenAI's real-world efficacy in enhancing operational performance. To address this gap, five GenAI platforms were deployed sequentially over five weeks with a nine-member team, applying identical prompts and project data. Key performance indicators for both supervisory activities (planning time, review load, and efficiency) and team outputs (volume processed, processing speed, and error rates) were recorded and compared against a baseline week. Quantitative analysis revealed improvements of up to 17% in throughput, reductions of up to 51% in error rates, and supervisory time savings of up to 47%. Qualitative surveys captured perceptions of clarity, relevance, and usability from both the project supervisor and team. Among the tools evaluated, ChatGPT Pro (GPT-4o) demonstrated the most balanced performance, excelling across multiple metrics and achieving the highest user satisfaction scores. The findings support a hybrid management model—combining AI-driven insights with human judgment—and outline best practices for prompt design, data governance, and continuous validation when integrating GenAI into project workflows.

ÍNDICE

RESUMEN.....	3
ABSTRACT.....	4
ÍNDICE	5
LISTADO DE TABLAS	7
LISTADO DE FIGURAS	9
CAPÍTULO I – DEFINICIÓN DEL PROYECTO DE TITULACIÓN	10
1.1 TEMA	10
1.2 JUSTIFICACIÓN	10
1.3 PLANTEAMIENTO DEL PROBLEMA.....	10
1.4 OBJETIVOS	11
1.5 ALCANCE DEL PROYECTO DE TITULACIÓN.....	12
1.6 METODOLOGÍA DE INVESTIGACIÓN	14
1.7 MARCO TEÓRICO.....	16
1.8 CRONOGRAMA.....	18
CAPÍTULO II – ANÁLISIS DE HERRAMIENTAS DE GEN-IA.....	19
2.1 INTRODUCCIÓN AL USO DE GENAI EN LA GESTIÓN DE PROYECTOS	19
2.1.1 El papel emergente de la GenAI en la gestión de proyectos.....	19
2.1.2 Beneficios observados en la gestión de proyectos mediante GenAI	19
2.1.3 Fases de la gestión de proyectos	20
2.2 CRITERIOS PARA LA SELECCIÓN DE HERRAMIENTAS	21
2.2.1 Metodología y criterios de selección	22
2.3 DESCRIPCIÓN GENERAL DE LAS HERRAMIENTAS ELEGIDAS	23
2.3.1 Presentación de las herramientas evaluadas.....	23
2.3.2 Características y funcionalidades de cada herramienta.....	26
2.3.3 Limitaciones identificadas en cada herramienta	29
2.3.4 Comparativa de las herramientas	31
CAPÍTULO III – EVALUACIÓN DEL CASO DE ESTUDIO	33
3.1 INTRODUCCIÓN AL CASO DE ESTUDIO	33
3.1.1 Descripción del proyecto	33
3.2 DISEÑO DEL ESTUDIO COMPARATIVO	37
3.2.1 Metodología de comparación y criterios de evaluación.....	37
3.3 PROTOCOLO DE IMPLEMENTACIÓN Y ESTANDARIZACIÓN DE	
INSTRUCCIONES	42

3.3.1 Esquema de implementación	42
3.3.2 Instrucciones proporcionadas a las herramientas GenAI.....	43
3.4 IMPLEMENTACIÓN Y EVALUACIÓN DE HERRAMIENTAS DE GENAI..	46
3.4.1 Resultados de la implementación con ChatGPT.....	46
3.4.2 Resultados de la implementación con Claude.....	50
3.4.3 Resultados de la implementación con Gemini.....	53
3.4.4 Resultados de la implementación con DeepSeek.....	58
3.4.5 Resultados de la implementación con Infinity.....	62
3.5 SÍNTESIS COMPARATIVA ENTRE HERRAMIENTAS.....	66
3.5.1 Tabla consolidada de desempeño por herramienta y KPI	66
CAPÍTULO IV – ANÁLISIS COMPARATIVO Y TOMA DE DECISIONES.....	68
4.1 RESUMEN DE LOS PRINCIPALES HALLAZGOS.....	68
4.1.1 Comparación cuantitativa de KPIs.....	68
4.1.2 Análisis cualitativo de las recomendaciones.....	79
4.2 PROPUESTA DE LA HERRAMIENTA MÁS ADECUADA PARA LA GESTIÓN DE PROYECTOS	81
4.2.1 Herramienta recomendada	81
4.3 IMPACTO DE LA APLICACIÓN DE GENAI EN LA GESTIÓN DE PROYECTOS 84	
4.3.1 Impacto en procesos y gestión	84
4.4 RECOMENDACIONES PARA LA APLICACIÓN DE GENAI EN LA GESTIÓN DE PROYECTOS	85
4.4.1 Buenas prácticas para el uso de GenAI en proyectos	85
4.4.2 Factores críticos de éxito.....	86
CONCLUSIONES Y RECOMENDACIONES.....	87
BIBLIOGRAFÍA	90
ANEXOS	93

LISTADO DE TABLAS

Tabla 1 Cronograma GenAI	18
Tabla 2 Comparación funcional de herramientas GenAI para gestión de proyectos	31
Tabla 3 Indicadores de desempeño semanal del supervisor del proyecto	39
Tabla 4 Indicadores operativos semanales del equipo de recolección y etiquetado	40
Tabla 5 Estimación de costos mensuales del proyecto	41
Tabla 6 Comparación de indicadores de desempeño del supervisor: semana base vs. uso de ChatGPT	48
Tabla 7 Comparación de indicadores de desempeño del equipo de recolección y etiquetado: Semana base vs. uso de ChatGPT	48
Tabla 8 Comparación de indicadores de desempeño del supervisor: semana base vs. uso de Claude.....	51
Tabla 9 Comparación de indicadores de desempeño del equipo de recolección y etiquetado: semana base vs. uso de Claude.....	52
Tabla 10 Comparación de indicadores de desempeño del supervisor: semana base vs. uso de Gemini.....	56
Tabla 11 Comparación de indicadores de desempeño del equipo de recolección y etiquetado: semana base vs. uso de Gemini	56
Tabla 12 Comparación de indicadores de desempeño del supervisor: semana base vs. uso de DeepSeek.....	60
Tabla 13 Comparación de indicadores de desempeño del equipo de recolección y etiquetado: semana base vs. uso de DeepSeek.....	60
Tabla 14 Comparación de indicadores de desempeño del supervisor: semana base vs. uso de Infinity	64
Tabla 15 Comparación de indicadores de desempeño del equipo de recolección y etiquetado: semana base vs. uso de Infinity	64
Tabla 16 Comparación de indicadores de desempeño del supervisor: semana base vs. herramientas de GenAI	66
Tabla 17 Comparación de indicadores de desempeño del equipo de recolección y etiquetado: semana base vs. herramientas de GenAI.....	67
Tabla 18 Valores absolutos de los KPIs clave de supervisión: semana base vs. herramientas de GenAI	68

Tabla 19 Valores absolutos de los KPIs clave del equipo: semana base vs. herramientas de GenAI 71

LISTADO DE FIGURAS

Figura 1 Flujo semanal del supervisor del proyecto	34
Figura 2 Flujo semanal del equipo de recolección y etiquetado	35
Figura 3 Variación porcentual del tiempo total de supervisión.....	69
Figura 4 Variación porcentual de la eficiencia de planificación	69
Figura 5 Variación porcentual del número de muestras revisadas	70
Figura 6 Variación porcentual del tiempo dedicado a revisión diaria.....	71
Figura 7 Variación porcentual del total de imágenes recolectadas	72
Figura 8 Variación porcentual de la velocidad de recolección.....	73
Figura 9 Variación porcentual de la tasa de error en recolección	73
Figura 10 Variación porcentual del total de imágenes etiquetadas	74
Figura 11 Variación porcentual de la velocidad de etiquetado.....	75
Figura 12 Variación porcentual de la tasa de error en etiquetado	75
Figura 13 Variación porcentual de la productividad general	76
Figura 14 Variación porcentual del tiempo total de subida.....	77
Figura 15 Variación porcentual del costo por imagen.....	77
Figura 16 Tiempo de respuesta de las herramientas GenAI.....	78
Figura 17 Puntuación media global del supervisor.....	79
Figura 18 Puntuación media global del equipo.....	80
Figura 19 Puntuación media en criterios técnicos por herramienta	81

CAPÍTULO I – DEFINICIÓN DEL PROYECTO DE TITULACIÓN

1.1 TEMA

Integración de herramientas de Inteligencia Artificial Generativa en la gestión de proyectos: un estudio aplicado

1.2 JUSTIFICACIÓN

La Inteligencia Artificial Generativa (GenAI) ha revolucionado diversas industrias, ofreciendo nuevas formas de optimizar procesos y mejorar la eficiencia operativa. En el ámbito de la gestión de proyectos, el uso de herramientas de GenAI promete transformar la manera en que se gestionan las tareas, los recursos y las decisiones. Estas herramientas, gracias a sus diversas funcionalidades orientadas a la eficiencia y al soporte en la gestión, tienen el potencial de optimizar diferentes aspectos del trabajo en proyectos. Sin embargo, a pesar de la creciente adopción de la IA, existe una falta de estudios prácticos que evalúen la efectividad real de las herramientas de GenAI en proyectos reales. Aunque el potencial teórico de estas tecnologías está ampliamente documentado, la evidencia empírica que demuestra su impacto en la gestión de proyectos reales sigue siendo limitada. Este estudio busca responder a esta cuestión evaluando diversas herramientas de Inteligencia Artificial Generativa aplicadas a un proyecto real. Mediante este análisis, se busca demostrar la utilidad y confiabilidad de estas herramientas, identificando aquella que ofrece la mayor eficiencia. Al finalizar, se propondrá la herramienta de GenAI más adecuada para la gestión de proyectos, con el respaldo de datos obtenidos en un entorno real.

1.3 PLANTEAMIENTO DEL PROBLEMA

Aunque la Inteligencia Artificial Generativa promete transformar la gestión de proyectos, mejorando la eficiencia y reduciendo los errores humanos, aún existe una falta de evidencia práctica que demuestre su efectividad en entornos reales. Las empresas, a pesar de estar interesadas en adoptar estas herramientas, se encuentran con la incertidumbre de si realmente aportarán valor en términos de productividad y calidad. Además, surge la pregunta de cuál herramienta de GenAI es la más adecuada. La escasez de estudios que validen el impacto de GenAI en proyectos reales y que comparen diferentes herramientas crea una brecha en el conocimiento, dificultando su implementación efectiva. Este estudio busca abordar esta problemática evaluando diversas herramientas de GenAI en un caso de estudio real, proporcionando información confiable sobre su utilidad en la gestión de proyectos y con el objetivo de identificar cuál de ellas ofrece el mayor valor en la gestión de proyectos.

Identificación del Problema Principal

El principal problema que aborda este estudio es la falta de evidencia práctica y comparativa sobre la efectividad de las herramientas de Inteligencia Artificial Generativa en la gestión de proyectos. En un entorno donde la precisión y la eficiencia son crucial, las empresas necesitan guías confiables que les permitan escoger las herramientas más adecuadas para sus proyectos. Este estudio se propone llenar esta brecha, evaluando varias herramientas de GenAI en un proyecto real para demostrar su utilidad, fiabilidad y aplicabilidad en el ámbito de la gestión de proyectos.

Problemas Secundarios

- Falta de estudios comparativos: No existen suficientes estudios que comparen las distintas herramientas de GenAI en contextos reales, lo que impide a las empresas elegir la opción más eficiente para sus proyectos.
- Desconfianza en la efectividad de GenAI: Las organizaciones tienen dudas sobre si las herramientas de GenAI realmente mejorarán sus procesos, debido a la escasez de datos concretos que demuestren su impacto en la gestión de proyectos.
- Falta de evidencia sobre el impacto en métricas clave de desempeño: La mayoría de los estudios existentes se centran en capacidades generales de la GenAI, pero no ofrecen datos concretos sobre cómo estas herramientas afectan indicadores operativos como la eficiencia o la productividad en la gestión de proyectos.

Formulación del Problema

De manera general, el problema puede formularse en la pregunta: ¿Qué herramientas de GenAI son más eficaces para la gestión de proyectos, y cómo puede demostrarse su confiabilidad y eficiencia en un entorno práctico? Para abordar esta cuestión, se analizarán y compararán diversas herramientas de GenAI, considerando métricas de eficiencia y un estudio de caso en un proyecto específico, con el fin de desarrollar un marco de decisión aplicable a gestión de proyectos.

1.4 OBJETIVOS

Objetivo General

Proponer un marco de evaluación y selección de herramientas de Inteligencia Artificial Generativa aplicadas a la gestión de proyectos, mediante un análisis comparativo y un estudio de caso práctico.

Objetivos Específicos

1. Analizar una selección estratégica de herramientas de GenAI aplicables a la gestión de proyectos, considerando su funcionalidad, enfoque de uso y accesibilidad para su evaluación comparativa.
2. Realizar un estudio comparativo de herramientas de GenAI, aplicando métricas de eficiencia en un caso de estudio real.
3. Formular recomendaciones fundamentales sobre la herramienta de GenAI más adecuada para la gestión de proyectos, basada en el análisis comparativo y los resultados del caso de estudio.
4. Presentar conclusiones sobre la efectividad y eficiencia del uso de herramientas de GenAI en la gestión de proyectos, y proponer una recomendación estratégica basada en los resultados obtenidos.

1.5 ALCANCE DEL PROYECTO DE TITULACIÓN

Este proyecto se centra en evaluar la efectividad y eficiencia del uso de herramientas de Inteligencia Artificial Generativa en la gestión de proyectos, mediante un análisis comparativo aplicado a un caso de estudio real. A partir de los resultados obtenidos, se formularán recomendaciones estratégicas sobre la herramienta de GenAI más adecuada, con el fin de guiar a las organizaciones en su adopción informada. El alcance del estudio incluye:

Evaluación de Herramientas de GenAI

- Se seleccionarán y evaluarán diversas herramientas de GenAI utilizadas en la gestión de proyectos, analizando sus funcionalidades clave, accesibilidad, y eficiencia en el contexto práctico.

Análisis Comparativo

- Se llevará a cabo un estudio comparativo de las herramientas de GenAI seleccionadas, utilizando métricas de eficiencia y desempeño, basadas en un caso de estudio real de gestión de proyectos.

Desarrollo de Recomendaciones

- El producto final será una recomendación detallada, respaldada por un marco de decisión, sobre la herramienta de GenAI más adecuada para la gestión de proyectos, basada en evidencia de su efectividad y confiabilidad en un entorno real.

Requisitos del Proyecto

Los requisitos de este proyecto han sido definidos para asegurar que el análisis y los resultados cumplan con los estándares de calidad y aplicabilidad esperados:

- **Requisitos Técnicos:** Las herramientas de IA evaluadas deben ofrecer funcionalidades avanzadas para la gestión de proyectos.
- **Requisitos de Evaluación:** Cada herramienta seleccionada será evaluada utilizando métricas de eficiencia y desempeño en un caso de estudio real.
- **Requisitos de Documentación:** El estudio incluirá la documentación completa de los resultados obtenidos, con un análisis detallado y la creación de un marco de decisión accesible y útil.

Consideraciones del Proyecto

El estudio se basa en un caso de uso real conformado por tareas operativas semanales, estructuradas y gestionadas como proyectos parciales con metas claras y procesos definidos. Aunque no abarca todas las fases formales de un proyecto completo, este enfoque permite evaluar de manera efectiva el rendimiento de herramientas de GenAI en un entorno repetible. La elección de este tipo de caso respondió a la necesidad de garantizar la viabilidad del estudio dentro del tiempo disponible para su desarrollo. Implementar un proyecto integral habría dificultado la aplicación equitativa de múltiples herramientas y la recolección sistemática de datos comparables. En cambio, el modelo adoptado facilita una evaluación objetiva, cuantificable y alineada con los objetivos de investigación.

Criterios de Aceptación

Para que el proyecto sea considerado exitoso, la guía final debe:

- Proporcionar evidencia confiable sobre la efectividad de las herramientas de GenAI en la gestión de proyectos reales.
- Incluir un análisis comparativo detallado de las herramientas evaluadas, con métricas de eficiencia y desempeño en el contexto del estudio de caso.
- Generar un marco de decisión claro y aplicable para la selección de herramientas de GenAI, basado en los resultados obtenidos.

Exclusiones

El presente estudio no contempla el diseño ni el desarrollo de modelos de Inteligencia Artificial personalizados, ni la implementación directa de las herramientas seleccionadas en los sistemas productivos de toda la empresa. El alcance de la investigación se limita a la evaluación comparativa y aplicación práctica de herramientas de GenAI ya existentes, en un entorno real. Aunque se trata de un estudio comparativo, su carácter es eminentemente aplicado, ya que las herramientas fueron implementadas de manera real en el flujo de trabajo del equipo, permitiendo recopilar datos concretos. Debido a la naturaleza confidencial del entorno empresarial, no se incluyen detalles específicos ni se presentan datos sensibles relacionados con la organización, el equipo de trabajo o los proyectos involucrados. Asimismo, el estudio no abarca la personalización ni el desarrollo técnico de las herramientas de GenAI evaluadas, ni su integración en plataformas tecnológicas existentes.

Restricciones

El análisis se limitará a herramientas de GenAI que ofrezcan versiones de pago o gratuitas, según la disponibilidad. Si están disponibles versiones premium, se utilizarán esas; si solo hay versiones gratuitas u open-source, se usará la opción correspondiente para la evaluación.

1.6 METODOLOGÍA DE INVESTIGACIÓN

La metodología adoptada en este estudio corresponde a una investigación aplicada de tipo tecnológico, con un enfoque experimental, cuantitativo y comparativo, orientada a evaluar el impacto real de herramientas de Inteligencia Artificial Generativa (GenAI) en la gestión de proyectos. Se busca no solo validar la utilidad y eficiencia de estas herramientas en un entorno práctico, sino también formular una recomendación fundamentada sobre su adopción.

Tipo de Investigación y Diseño

El diseño metodológico está basado en un estudio de caso, aplicado a un proyecto real y recurrente de recolección y etiquetado de imágenes.

Descripción del Caso de Estudio

El caso de estudio utilizado para esta investigación corresponde a un proyecto semanal de recolección y etiquetado de imágenes, gestionado por un equipo de nueve miembros. El propósito principal de esta recolección es generar conjuntos de datos visuales de alta calidad para el entrenamiento de modelos de inteligencia artificial, específicamente en tareas de visión por

computadora. Cada semana, el equipo tiene como objetivo recolectar y etiquetar aproximadamente 2,500 imágenes, cumpliendo metas diarias individuales bajo estándares previamente definidos de calidad y productividad. La naturaleza colaborativa, repetitiva y medible de este proyecto lo convierte en un entorno ideal para experimentar con herramientas de GenAI y observar su impacto en la eficiencia, la gestión del tiempo y los costos asociados.

Fases de la Investigación

1. **Aplicación en el Caso de Estudio:** A lo largo de cinco semanas consecutivas, dentro del proyecto, se aplicará una herramienta de GenAI diferente cada semana. Cada herramienta será utilizada para analizar la documentación y los datos operativos ya recolectados del equipo y del proyecto. A partir de este análisis, se solicitarán recomendaciones específicas orientadas a mejorar la gestión del proyecto, la productividad del equipo y otros aspectos claves. Estas recomendaciones serán generadas a partir de un prompt detallado enviado a cada herramienta de GenAI, el cual incluirá información operativa del proyecto, métricas clave de desempeño y objetivos de mejora. Las recomendaciones generadas por cada herramienta serán implementadas durante esa misma semana, lo que permitirá observar y medir su impacto real en un entorno operativo.
2. **Recopilación y Registro de Datos:** Durante cada semana, se recopilarán datos relacionados con el rendimiento del equipo. Se registrarán tiempos de ejecución, volumen de imágenes procesadas, errores detectados, y demás datos clave. Esta información será organizada y analizada con base en indicadores clave de desempeño previamente definidos.
3. **Evaluación Comparativa:** Con base en los datos recolectados, se realizará una comparación objetiva del desempeño de las herramientas, utilizando métricas cuantificables. Esto permitirá determinar con claridad cuál herramienta ofrece mejores resultados.
4. **Análisis Final y Recomendaciones:** A partir del análisis, se generará un marco de decisión que identifique la herramienta más adecuada según los resultados obtenidos en el caso de estudio, integrando tanto el rendimiento cuantitativo como la experiencia práctica del equipo durante la implementación.

Técnicas de Análisis

Se utilizarán instrumentos de medición y análisis tanto cuantitativos como cualitativos, incluyendo:

- Hojas de cálculo de seguimiento diario para recolección de datos operativos (tiempo, errores, imágenes procesadas).
- Encuestas semanales al equipo y al supervisor para evaluar la percepción sobre la organización de tareas, y en el caso del supervisor, la utilidad y aplicabilidad de cada herramienta GenAI.
- Técnicas estadísticas básicas para análisis comparativo de KPIs.
- Análisis de tendencias semanales para observar mejoras.

Aunque el estudio tiene un enfoque comparativo, su carácter es principalmente aplicado, ya que cada herramienta será implementada de forma práctica en el entorno laboral del equipo, permitiendo observar su comportamiento en situaciones reales y operativas.

Consideraciones Éticas

Aunque el estudio se desarrolla en un entorno real, se han tomado todas las medidas necesarias para proteger la confidencialidad de los datos. No se incluirán nombres de la empresa, proyectos específicos ni información sensible. El experimento cuenta con la autorización de la empresa, pero no forma parte de un proceso oficial corporativo, sino de un estudio con fines académicos.

1.7 MARCO TEÓRICO

Inteligencia Artificial (IA)

La Inteligencia Artificial (IA) se refiere a la capacidad de las máquinas y sistemas informáticos para realizar tareas que tradicionalmente requieren de la inteligencia humana, como el aprendizaje, la toma de decisiones, el reconocimiento de patrones y el procesamiento de lenguaje natural. En el contexto actual, la IA es fundamental para diversas aplicaciones tecnológicas, desde los asistentes virtuales hasta los sistemas de diagnóstico médico, y está transformando industrias como la automotriz, la salud y las finanzas. Su evolución ha sido rápida, pasando de sistemas básicos de automatización en los años 50 a sofisticados algoritmos de aprendizaje profundo que impulsan la IA moderna (IBM, 2024).

Inteligencia Artificial Generativa (GenAI)

La Inteligencia Artificial Generativa (GenAI) es un tipo de IA que tiene la capacidad de generar nuevos contenidos, soluciones o predicciones a partir de datos de entrada. A diferencia de las IA tradicionales, que se centran en la clasificación o el análisis de datos existentes, GenAI crea resultados originales, como imágenes, texto o música. Su principal característica es la capacidad de “crear”, lo que la hace útil en áreas como la creación de contenido automatizada, el diseño de productos y la mejora de procesos de negocio. GenAI ha ganado tracción en diversas industrias, como la creación de contenido en medios, la medicina personalizada, y la automatización de tareas de diseño. Las aplicaciones de GenAI dentro de las organizaciones incluyen la automatización de procesos repetitivos, la mejora en la atención al cliente, la creación de contenido personalizado, y la optimización de la toma de decisiones. GenAI permite a las empresas adaptarse rápidamente a los cambios del mercado y mejorar su competitividad (McKinsey & Company, 2023).

Gestión de Proyectos

La gestión de proyectos es el proceso de planificar, organizar, ejecutar, controlar y finalizar un proyecto de manera eficiente, alcanzando los objetivos establecidos dentro de los plazos y presupuesto definidos. Abarca cinco fases clave: inicio, planificación, ejecución, monitoreo y cierre. La gestión de proyectos es crucial para el éxito organizacional, ya que asegura que los recursos se utilicen de manera efectiva, se gestionen los riesgos y se entreguen los proyectos a tiempo y dentro del presupuesto. Mientras que las metodologías tradicionales, como el modelo Waterfall, se enfocan en procesos lineales, enfoques más modernos como Agile o Scrum permiten adaptarse mejor a cambios rápidos y a necesidades emergentes durante el ciclo de vida del proyecto (Project Management Institute, 2017).

Análisis Comparativo

El análisis comparativo implica la evaluación de diferentes opciones, herramientas o enfoques para determinar cuál ofrece el mayor valor o rendimiento en función de criterios específicos. En el contexto de este estudio, se llevará a cabo un análisis comparativo de diversas herramientas de Inteligencia Artificial Generativa (GenAI) aplicadas a la gestión de proyectos, utilizando métricas de eficiencia, facilidad de uso, impacto en la toma de decisiones y la mejora de procesos operativos. Este tipo de análisis es esencial para identificar cuál de las herramientas evaluadas tiene el mayor potencial para optimizar la gestión de proyectos (Rodríguez & Bastidas, 2013).

Automatización y Optimización de Procesos

La automatización de procesos mediante IA implica el uso de herramientas tecnológicas para realizar tareas repetitivas de manera eficiente, reduciendo la intervención humana y minimizando errores. La optimización de procesos va un paso más allá, buscando mejorar la eficiencia y efectividad de las operaciones al eliminar ineficiencias y hacer un mejor uso de los recursos. En la gestión de proyectos, estas tecnologías permiten automatizar tareas como la programación de actividades y la asignación de recursos, lo que resulta en una mayor rapidez en la ejecución del proyecto y una mejor toma de decisiones (Parra Mora & Bolívar Cazañas, 2018).

Integración de Tecnología en la Gestión de Proyectos

La incorporación de tecnologías emergentes en la gestión de proyectos está redefiniendo la forma en que los equipos planifican, ejecutan y monitorean sus actividades. Esta transformación permite una mayor agilidad operativa, mejora la eficiencia en la asignación de recursos y facilita la adaptabilidad frente a cambios imprevistos. Además, la tecnología no solo contribuye a optimizar la ejecución de tareas, sino que también fortalece la toma de decisiones mediante el acceso a datos en tiempo real y la capacidad de anticipar escenarios futuros (Project Management Institute, 2017).

1.8 CRONOGRAMA

Tabla 1

Cronograma GenAI

Actividad	Duración	Fecha de inicio	Fecha de finalización
Revisión bibliográfica y recopilación de información	2 semanas	01/01/2025	15/01/2025
Selección y análisis de herramientas de GenAI	1 semana	15/01/2025	22/01/2025
Preparar del caso de estudio	2 semanas	22/01/2025	05/02/2025
Implementación práctica en el caso de estudio	5 semanas	05/02/2025	12/03/2025
Evaluación de resultados	2 semanas	12/03/2025	26/03/2025
Desarrollo del marco de decisión y recomendaciones	2 semanas	26/03/2025	09/04/2025
Redacción del informe final	4 semanas	09/04/2025	07/05/2025
Revisión y ajustes finales	2 semanas	07/05/2025	21/05/2025

Nota. El cronograma muestra de forma clara y secuencial las ocho fases del proyecto, indicando su duración y fechas exactas, lo que facilita el seguimiento y control de cada hito.

CAPÍTULO II – ANÁLISIS DE HERRAMIENTAS DE GEN-IA

2.1 INTRODUCCIÓN AL USO DE GENAI EN LA GESTIÓN DE PROYECTOS

En la última década, la gestión de proyectos ha experimentado una transformación profunda impulsada por tecnologías emergentes. Entre ellas, la Inteligencia Artificial Generativa (GenAI) ha destacado por su capacidad para ofrecer nuevas soluciones en áreas clave como la planificación, la toma de decisiones y el análisis de datos.

2.1.1 El papel emergente de la GenAI en la gestión de proyectos

En los últimos años, la Inteligencia Artificial Generativa (GenAI) ha dejado de ser una promesa tecnológica emergente para convertirse en una herramienta concreta con aplicaciones relevantes en múltiples sectores, incluyendo la gestión de proyectos. A diferencia de otros tipos de inteligencia artificial centrados en el análisis de datos históricos o la automatización de tareas específicas, la GenAI introduce la capacidad de generar nuevas ideas, soluciones y recomendaciones personalizadas a partir de grandes volúmenes de información. Esta característica ha despertado un creciente interés en su posible aplicación en la toma de decisiones estratégicas, la planificación de recursos y la mejora continua dentro de entornos de proyectos (McKinsey & Company, 2023). De acuerdo con Project-Management.com (2024), cada vez más empresas están incorporando soluciones basadas en GenAI para optimizar la planificación, anticiparse a desviaciones y adaptar los planes de trabajo en tiempo real, lo cual representa un cambio sustancial frente a los enfoques tradicionales documentados en el PMBOK® Guide (Project Management Institute, 2017). En conjunto, los hallazgos actuales demuestran que la GenAI posee un potencial transformador en la gestión de proyectos.

2.1.2 Beneficios observados en la gestión de proyectos mediante GenAI

La incorporación de herramientas basadas en Inteligencia Artificial Generativa (GenAI) en entornos de gestión de proyectos ha generado beneficios significativos, tanto en términos de eficiencia operativa como en la calidad de la toma de decisiones. Una de las principales ventajas observadas es la capacidad de estas tecnologías para analizar rápidamente grandes volúmenes de información y entregar recomendaciones en tiempo real, lo que permite a los equipos actuar con mayor agilidad y precisión (Project Management Institute, 2025). Además, el Project Management Institute (PMI) ha reconocido oficialmente el papel de la inteligencia artificial —y

en particular de la GenAI— como uno de los factores más influyentes en el futuro de la gestión de proyectos.

De acuerdo con el curso “Generative AI Overview for Project Managers” ofrecido por el PMI desde principios de 2025, la integración de GenAI en la práctica de la gestión de proyectos permite mejorar la planificación, ejecución e innovación en los proyectos. Las herramientas basadas en esta tecnología impulsan la productividad, la creatividad y la colaboración, mejoran la entrega de resultados y abren nuevas oportunidades profesionales para los líderes de proyecto. Asimismo, el Foro Económico Mundial proyecta que, para el año 2027, el 75 % de las empresas habrán adoptado tecnologías como la GenAI, impulsadas por la accesibilidad pública de modelos de lenguaje avanzados.

El artículo “First-Mover Advantage” del PMI refuerza esta tendencia al mostrar que el 99 % de los equipos pioneros (“trailblazers”) que adoptaron tempranamente GenAI han reportado resultados positivos. Más del 85 % de estos equipos destacan mejoras sustanciales en áreas clave como el alcance, cronograma, costo y gestión de la calidad del proyecto. También se evidencian avances importantes en la resolución de problemas, la efectividad del equipo, la innovación y la capacidad de colaboración (Project Management Institute, 2025).

Estos beneficios, aunque prometedores, deben ser evaluados cuidadosamente según el contexto de cada proyecto. Por ello, el presente estudio se enfoca en analizar comparativamente el desempeño real de varias herramientas de GenAI aplicadas a un mismo entorno, con el objetivo de validar su efectividad y formular una recomendación basada en evidencia.

2.1.3 Fases de la gestión de proyectos

La gestión de proyectos se articula en cinco fases interrelacionadas que aseguran el cumplimiento de objetivos y el uso eficiente de recursos:

- **Inicio:** Se identifican y documentan la visión, los objetivos y los interesados clave. Incluye la elaboración del acta de constitución, que formaliza el proyecto y define sus límites, requisitos y criterios de éxito.
- **Planificación:** Se desarrollan el plan de gestión del alcance, el cronograma de actividades, el presupuesto y la asignación de recursos. Aquí se detallan las tareas, se establecen hitos, se calculan costos, se evalúan riesgos y se diseñan los procedimientos de comunicación y control de calidad.

- **Ejecución:** Consiste en coordinar personas y recursos para llevar a cabo las actividades definidas. Se implementan los planes de trabajo, se gestionan los equipos y se generan los entregables. La ejecución demanda un seguimiento constante del progreso y la motivación del equipo para cumplir plazos y estándares.
- **Monitoreo y control:** Implica la medición y el seguimiento de indicadores clave (tiempos, costos, calidad), la identificación de desviaciones y la aplicación de acciones correctivas. Esta fase asegura que el proyecto se mantenga alineado con el plan original y permite la toma de decisiones informadas ante riesgos o cambios.
- **Cierre:** Culmina con la validación de los entregables, la aceptación formal por parte de los interesados y la liberación de recursos. Se realiza el informe final de lecciones aprendidas y se documenta toda la información para futuras referencias, garantizando la captura de buenas prácticas y oportunidades de mejora.

2.2 CRITERIOS PARA LA SELECCIÓN DE HERRAMIENTAS

La selección de las herramientas de inteligencia artificial generativa consideradas en esta investigación se realizó de manera estratégica, con base en criterios diseñados para asegurar su pertinencia, aplicabilidad y diversidad funcional dentro del contexto de la gestión de proyectos. Si bien no se llevó a cabo una revisión exhaustiva del mercado, se optó por conformar una muestra representativa, integrando herramientas que han sido destacadas por su uso frecuente en entornos profesionales y por su potencial para apoyar procesos clave en la dirección de proyectos. La elección se fundamentó en la revisión de literatura académica reciente sobre adopción tecnológica en entornos organizacionales (Kolbjørnsrud & Thomas, 2023; Project-Management.com, 2024) como las mejores prácticas observadas en el uso real de GenAI en la industria. Se procuró también incluir tanto asistentes conversacionales de propósito general como plataformas especializadas en gestión de proyectos, con el fin de evaluar enfoques diversos y comparar su efectividad bajo un marco homogéneo de análisis. Este proceso de selección tuvo como objetivo construir una muestra representativa y funcional que permitiera realizar un análisis comparativo riguroso, aplicable a casos reales y orientado a medir el impacto concreto de estas tecnologías en la eficiencia y productividad dentro de la gestión de proyectos. Cabe aclarar que el propósito de esta investigación no fue realizar una revisión exhaustiva del mercado de herramientas de GenAI, sino seleccionar una muestra funcional y representativa, adecuada para su aplicación práctica en un entorno real de gestión de proyectos.

2.2.1 Metodología y criterios de selección

La metodología empleada para la selección de herramientas de Inteligencia Artificial Generativa se diseñó con un enfoque práctico, estructurado y alineado con marcos conceptuales aceptados en estudios sobre adopción tecnológica (Venkatesh et al., 2003; Kolbjørnsrud & Thomas, 2023). Se aplicó una lógica comparativa funcional, apropiada para investigaciones exploratorias-aplicadas, donde el objetivo no es realizar un levantamiento exhaustivo del mercado, sino evaluar soluciones tecnológicas representativas en contextos reales. El proceso se desarrolló en tres etapas principales:

1. Exploración dirigida y preselección inicial:

Se identificaron herramientas ampliamente reconocidas por su uso activo en entornos profesionales y académicos, considerando factores como disponibilidad, accesibilidad, variedad funcional y popularidad. La selección inicial se orientó hacia plataformas que ya presentaban aplicaciones relevantes en actividades como planificación, seguimiento, automatización y apoyo a la toma de decisiones (Wells, 2023; HubSpot, 2024).

2. Aplicación de criterios básicos de evaluación:

En lugar de utilizar filtros teóricos genéricos, se priorizaron criterios prácticos enfocados en la aplicabilidad funcional y operativa de las herramientas en un entorno de gestión de proyectos. Los principales criterios considerados fueron:

- Representatividad funcional: inclusión de herramientas con enfoques generales y otras más especializadas, para abarcar distintos casos de uso.
- Accesibilidad y disponibilidad de funcionalidades avanzadas: preferencia por versiones completas (Pro o comerciales) que permitieran evaluar todo el potencial de cada herramienta.
- Facilidad de integración operativa: consideración de soluciones que pudieran aplicarse directamente, sin requerimientos técnicos complejos.
- Diversidad tecnológica: selección de herramientas provenientes de distintos proveedores, con arquitecturas y estrategias variadas, que facilitaran una comparación realista y equilibrada.

3. Selección final y alineación con el caso de estudio:

La muestra definitiva se conformó con cinco herramientas que cumplieran los criterios establecidos y que reflejaban distintas aproximaciones al uso de la GenAI en la gestión

de proyectos. Estas fueron seleccionadas no solo por su relevancia tecnológica, sino también por su viabilidad práctica de uso en el entorno específico del estudio. Se garantizó el acceso a versiones completas en todos los casos posibles, y en aquellas plataformas donde solo se disponía de versiones gratuitas, se evaluaron sin restricciones dentro de sus capacidades disponibles.

Esta metodología permitió conformar un conjunto de herramientas diverso, funcional y representativo, que sirviera como base para un análisis comparativo riguroso sobre el impacto de la GenAI en la eficiencia, productividad y toma de decisiones en contextos reales de gestión de proyectos.

2.3 DESCRIPCIÓN GENERAL DE LAS HERRAMIENTAS ELEGIDAS

2.3.1 Presentación de las herramientas evaluadas

A continuación, se describen las cinco herramientas de GenAI seleccionadas para este estudio, con sus versiones más recientes disponibles hasta mayo de 2025:

1. ChatGPT – Versión Pro con GPT-4o

- Desarrollador: OpenAI
- Fecha de lanzamiento: 13 de mayo de 2024
- Sobre la herramienta y la empresa: ChatGPT es el modelo de lenguaje más reconocido de OpenAI, una organización fundada en 2015 con el objetivo de desarrollar inteligencia artificial de manera segura y beneficiosa para la humanidad. La empresa es responsable de avances clave en el campo de los modelos de lenguaje natural. La empresa ha recibido inversiones significativas de Microsoft, acumulando más de \$10 mil millones en financiamiento. OpenAI ha desarrollado una línea de modelos desde GPT-1 (2018) hasta GPT-4o (2024), siendo este último multimodal (texto, imagen, audio) y altamente eficiente (OpenAI, 2024).
- GPT-4o (la “o” viene de “omni”) es la versión más reciente del modelo y representa un salto importante en rendimiento y versatilidad. Ofrece capacidades multimodales (texto, imágenes, audio) y ha sido optimizado para velocidad y eficiencia. En su versión Pro, permite generación de contenido, análisis de documentos, redacción estructurada, asistencia para codificación, creación de reportes, y mucho más, consolidándose como una herramienta integral para

profesionales de múltiples sectores, incluidos los que gestionan proyectos (OpenAI, 2024).

2. Claude – Claude 4 Sonnet

- Desarrollador: Anthropic
- Fecha de lanzamiento: 24 de febrero de 2025
- Sobre la herramienta y la empresa: Anthropic es una empresa fundada en 2021 por exmiembros de OpenAI, con una fuerte orientación hacia la investigación en IA segura y centrada en principios éticos. Su modelo de desarrollo se basa en principios de “IA constitucional”. Claude 4 Sonnet integra mejoras en razonamiento y manejo de texto complejo. Anthropic ha recibido inversiones destacadas de Google y Amazon, consolidándose como un actor influyente en la IA constitucional (Anthropic, 2025; AboutAmazon, 2025).
- Claude, su principal asistente de IA, ha sido nombrado en honor al pionero de la computación Claude Shannon. La versión 4 Sonnet se destaca por su capacidad para manejar documentos largos, razonar con mayor profundidad, y ofrecer respuestas equilibradas y bien argumentadas. Su uso es especialmente valorado en tareas que requieren precisión conceptual, redacción técnica, planificación y análisis estructurado. Su enfoque en seguridad lo ha posicionado entre las herramientas más confiables para usos profesionales (Anthropic, 2025).

3. Gemini – Gemini 2.5 Pro

- Desarrollador: Google DeepMind
- Fecha de lanzamiento: 25 de marzo de 2025
- Sobre la herramienta y la empresa: Gemini es la evolución de Bard, desarrollado por Google a través de su división de IA avanzada, DeepMind, fundada en 2010 y adquirida por Google en 2014, ha liderado proyectos como AlphaGo y AlphaFold. Gemini 2.5 Pro es multimodal (texto, imagen, audio, video) y está profundamente integrado con Google Workspace, lo que permite automatizar tareas directamente desde Google Docs, Sheets y Gmail. Esta integración lo convierte en una herramienta versátil para entornos colaborativos digitales (Google DeepMind, 2025; TechTarget, 2025).
- Gemini 2.5 Pro es la versión más potente de esta línea de modelos y se distingue por su integración nativa con Google Workspace. Esto permite usar la IA directamente dentro de documentos, hojas de cálculo, correos y presentaciones, lo

que facilita la automatización de tareas administrativas, la generación de contenido y la colaboración en tiempo real. Además, incorpora capacidades mejoradas para el análisis de datos, imágenes y video, siendo una herramienta versátil para cualquier entorno digital productivo (Google DeepMind, 2025).

4. DeepSeek – DeepSeek-V3-0324

- Desarrollador: DeepSeek
- Fecha de lanzamiento: 25 de marzo de 2025
- Sobre la herramienta y la empresa: DeepSeek es un startup tecnológica china apoyada por el fondo High-Flyer que se ha enfocado en el desarrollo de modelos de lenguaje de código abierto de alto rendimiento. Aunque es menos conocida en occidente, ha ganado relevancia en la comunidad de código abierto por ofrecer modelos con acceso gratuito y capacidades técnicas avanzadas. DeepSeek-V3-0324 cuenta con 671 mil millones de parámetros, siendo uno de los modelos más robustos de su tipo. El modelo fue entrenado con una inversión inicial reportada de \$5.5 millones, aunque estimaciones independientes sitúan el costo total por encima de \$1.3 mil millones (DeepSeek-VL Team, 2024).
- El modelo DeepSeek-V3-0324 es uno de los más robustos de su categoría, con más de 600 mil millones de parámetros. Su enfoque está centrado en precisión lógica, generación estructurada de texto y procesamiento de información técnica. Es ideal para tareas que requieren redacción formal, análisis exhaustivo, explicaciones detalladas y síntesis informativa (DataCamp, 2025). Su modelo de acceso abierto también lo convierte en una alternativa estratégica para organizaciones que buscan autonomía tecnológica.

5. Infinity AI – Versión 1.3.7

- Desarrollador: Hammoq
- Fecha de actualización: 1 de mayo de 2025
- Sobre la herramienta y la empresa: Hammoq es una empresa emergente con sede en EE.UU. enfocada en proveer soluciones de productividad mediante automatización inteligente, especialmente para pequeñas y medianas empresas. Infinity AI se desarrolló sobre una infraestructura previa de modelos open-source, adaptándolos a tareas operativas con una interfaz sencilla. Aunque no compite en profundidad analítica con modelos como GPT o Claude, su facilidad de uso y

orientación práctica la hacen útil para equipos que priorizan rapidez organizativa y ejecución diaria (Hammoq, 2025; TechCrunch, 2025).

- Infinity AI combina funciones de planificación, escritura automática, asistencia contextual y generación de tareas en una sola interfaz. Su enfoque en usabilidad lo convierte en una opción atractiva para usuarios que priorizan simplicidad, velocidad y resultados inmediatos (Hammoq, 2025). Aunque su capacidad de razonamiento profundo es limitada frente a modelos como GPT o Claude, su diseño ágil y orientado a la acción la posiciona como una herramienta útil en proyectos que demandan organización, síntesis rápida y coordinación eficiente.

2.3.2 Características y funcionalidades de cada herramienta

A continuación, se describen, de forma detallada y técnica, las características y funcionalidades específicas de cada una de las herramientas GenAI seleccionadas.

1. ChatGPT

ChatGPT Pro con GPT-4o ofrece una combinación robusta de funcionalidades orientadas tanto a usuarios técnicos como no técnicos. Su capacidad multimodal permite interpretar imágenes, texto y archivos PDF, lo que facilita el análisis documental automatizado. Admite la carga de múltiples archivos a la vez, que pueden ser comparados, resumidos o transformados en reportes con lenguajes formales (OpenAI, 2024). Entre sus funciones avanzadas destacan:

- Análisis de datos estructurados (por ejemplo, CSV y tablas): convierte hojas de cálculo en reportes, listas de tareas o cronogramas.
- Generación de contenido largo, como actas de reunión, informes ejecutivos y planes de proyecto, con segmentación lógica por secciones.
- Síntesis de múltiples fuentes: puede integrar información de varios archivos para generar una versión unificada, útil para informes semanales o documentación final.
- Soporte en codificación básica, especialmente útil si se necesita automatizar pequeñas tareas de procesamiento de datos en Python o generar fragmentos para dashboards.
- ChatGPT también permite exportar texto en formato Markdown, HTML o texto plano, lo que facilita su incorporación en otros sistemas. Su enfoque en precisión

contextual mejora la coherencia de entregables en proyectos largos o con múltiples fases.

2. Claude

Claude 4 Sonnet se caracteriza por su impresionante capacidad para manejar y analizar grandes volúmenes de texto (hasta 200,000 tokens por conversación). Esto permite cargar documentos extensos como propuestas técnicas, contratos, cronogramas o manuales, para analizarlos, resumirlos o transformarlos en nuevas versiones más comprensibles (Anthropic, 2025). Funcionalidades clave para proyectos incluyen:

- Reescritura y revisión crítica de textos, con recomendaciones de mejora por claridad, concisión o tono.
- Simulación de roles para toma de decisiones: por ejemplo, actuar como gerente de proyecto, líder técnico o cliente para ofrecer perspectivas específicas sobre un plan.
- Generación de contenido con estructura lógica compleja, incluyendo numeración automática, formato formal y desarrollo temático fluido (ideal para planes de acción).
- Evaluación de calidad textual y consistencia narrativa, útil para revisar entregables redactados por diferentes personas.
- Claude es menos versátil que ChatGPT en procesamiento de imágenes o archivos, pero su fuerza está en la profundidad de análisis y en mantener coherencia semántica en textos largos o críticos.

3. Gemini

Gemini 2.5 Pro se distingue por su integración avanzada con Google Workspace, lo que permite automatizar procesos documentales dentro de Google Docs, Sheets y Gmail sin necesidad de cambiar de entorno. Al estar completamente interconectado con Google Drive, puede acceder a archivos previamente cargados, analizarlos y generar entregables directamente desde ahí (Google DeepMind, 2025). Funciones destacadas para la gestión de proyectos incluyen:

- Auto-resumen de documentos largos con formato sugerido (bullet points, tablas, encabezados).
- Asistencia para redacción colaborativa en tiempo real en Docs.

- Extracción de datos desde hojas de cálculo, con generación de gráficas y descripciones interpretativas.
- Generación de presentaciones (Slides) a partir de documentos o datos operativos, ideal para reuniones de avance.
- Su modo de trabajo es natural para equipos que ya utilizan el ecosistema Google, y permite crear, revisar y entregar documentación sin salir de un único entorno, lo que reduce fricción operativa y aumenta la eficiencia.

4. DeepSeek

DeepSeek-V3-0324, aunque menos conocido, destaca por su rendimiento técnico en tareas de redacción estructurada, razonamiento lógico y precisión en lenguaje técnico. Su modelo de gran tamaño permite sostener argumentos detallados y responder a preguntas especializadas con profundidad, lo que lo hace ideal para entornos donde la exactitud del lenguaje es crítica (DataCamp, 2025). Características específicas:

- Síntesis técnica y científica: redacta descripciones precisas sobre procesos o procedimientos operativos.
- Formato formal automatizado, como informes técnicos, plantillas de documentación, listas de verificación y pasos secuenciales.
- Resolución de problemas complejos usando lógica simbólica o deductiva.
- Soporte avanzado en pseudocódigo y explicaciones algorítmicas, útil para informes de ingeniería, auditorías técnicas o reportes con implicaciones matemáticas.
- DeepSeek no está diseñado para tareas administrativas o generación ligera de contenido, pero su fortaleza está en casos donde se necesita lenguaje estructurado, especializado y riguroso.

5. Infinity AI

Infinity AI 1.3.7 está diseñado para facilitar tareas diarias y mejorar la organización sin una curva de aprendizaje elevada. Aunque no posee capacidades avanzadas de análisis o procesamiento de datos complejos, ofrece agilidad en la gestión de actividades operativas y documentación simple (Hammoq, 2025). Sus funcionalidades más relevantes incluyen:

- Generación automática de listas de tareas, con subtareas, prioridades y fechas límite.

- Creación rápida de mensajes de seguimiento, emails, actualizaciones de estado o avisos para el equipo.
- Redacción de minutas y agendas de reuniones con plantillas simples.
- Planificación visual mediante flujos de trabajo, ideales para seguimiento de tareas individuales o equipos pequeños.
- Es ideal como complemento organizacional en proyectos donde la velocidad y simplicidad son más importantes que la profundidad analítica. También sirve como apoyo para roles operativos que necesitan generar textos rápidos con un lenguaje claro y directo.

2.3.3 Limitaciones identificadas en cada herramienta

Durante el análisis práctico de las herramientas GenAI seleccionadas, se identificaron diversas limitaciones específicas que pueden influir en su utilidad y aplicabilidad en el contexto de la gestión de proyectos. Estas limitaciones no necesariamente descalifican su uso, pero sí deben considerarse al momento de integrarlas en flujos de trabajo reales.

1. ChatGPT

A pesar de su versatilidad (OpenAI, 2024), ChatGPT presenta ciertas restricciones funcionales relevantes:

- Tendencia a sobre explicar o extenderse en exceso: al no tener delimitaciones claras, puede generar respuestas demasiado largas o redundantes, lo cual afecta la agilidad en tareas puntuales.
- Necesidad de indicaciones precisas: su rendimiento disminuye si no se le proporcionan instrucciones bien estructuradas; en tareas poco específicas puede generar resultados genéricos o fuera de contexto.
- Dificultad para mantener consistencia entre sesiones: aunque es capaz de recordar información dentro de una conversación, no mantiene memoria entre sesiones a menos que se use la función específica, lo cual limita el seguimiento continuo de un mismo proyecto.

2. Claude

Claude se destaca por su capacidad de análisis textual, pero tiene limitaciones que reducen su aplicabilidad en ciertos entornos (Anthropic, 2025):

- Falta de soporte para archivos no textuales: no puede procesar imágenes, hojas de cálculo ni otros formatos más allá del texto plano, lo cual lo limita en tareas de análisis de datos o contenido visual.
- Menor velocidad en respuestas extensas: cuando se le solicita desarrollar documentos largos con estructura compleja, su tiempo de respuesta puede ser mayor que el de otras herramientas.
- Sin integración directa con plataformas de productividad: a diferencia de Gemini o ChatGPT, no cuenta con conectores nativos para entornos como Google Workspace o gestores de tareas, lo que obliga a usarlo de forma aislada.

3. Gemini

Aunque destaca por su integración (Google DeepMind, 2025), Gemini presenta algunas restricciones técnicas y operativas:

- Limitaciones en tareas altamente técnicas: su capacidad para generar contenido técnico profundo o con estructura muy precisa es inferior a la de modelos como Claude o DeepSeek.
- Dependencia operativa del ecosistema Google: muchas de sus funciones más útiles solo se activan al utilizar Docs, Sheets u otras aplicaciones de Google, lo que lo hace menos flexible para equipos que utilizan otras plataformas.
- Capacidad contextual más limitada: en tareas que requieren mantener coherencia en bloques largos de información, tiende a fragmentar ideas o simplificar en exceso los conceptos desarrollados.

4. DeepSeek

Este modelo demuestra gran capacidad técnica (DataCamp, 2025), pero su aplicación práctica está sujeta a varios condicionantes:

- Interfaz y entorno poco amigable para usuarios no técnicos: su uso requiere familiaridad con herramientas de desarrollador o entornos de texto plano, lo que reduce su accesibilidad para usuarios operativos.
- Ausencia de herramientas de apoyo para tareas administrativas: no cuenta con funciones específicas para gestión de proyectos, como creación de cronogramas, reportes ejecutivos o automatización de tareas repetitivas.

- Curva de aprendizaje elevada para aprovechar su potencial: para obtener respuestas útiles y especializadas, se requiere formular indicaciones muy técnicas, lo que puede no ser viable para todos los perfiles dentro de un equipo de proyecto.

5. Infinity AI

Infinity AI está orientada a la simplicidad y productividad ágil (Hammoq, 2025), pero esto implica limitaciones importantes:

- Capacidad limitada para análisis profundo: no está diseñada para desarrollar contenido extenso, ni para tareas de razonamiento avanzado o escritura técnica.
- Falta de soporte para entradas complejas o múltiples: no permite trabajar con varios documentos a la vez, ni analizar estructuras jerárquicas de información.
- Ausencia de personalización avanzada o configuraciones específicas: su estructura es fija y poco adaptable, lo que limita su uso en proyectos que requieren plantillas, formatos institucionales o flujos personalizados.

2.3.4 Comparativa de las herramientas

Tabla 2

Comparación funcional de herramientas GenAI para gestión de proyectos

Herramienta	Versión evaluada	Acceso y modelo de uso	Multimodalidad	Ventajas principales	Limitaciones clave	Soporte para tareas complejas
ChatGPT	GPT-4o	Pro (pago mensual individual de \$20 USD)	Texto, imágenes, audio, archivos PDF	Multimodal, versátil, potente para análisis documental	Requiere instrucciones precisas, sin memoria entre sesiones	Alto - razonamiento, análisis de datos, planificación estructurada
Claude	Claude 4 Sonnet	Pro (pago mensual individual de \$17 USD)	Solo texto	Análisis profundo, coherencia en textos largos	Sin soporte para archivos no textuales, sin integraciones	Alto - redacción lógica, simulación de escenarios, revisión técnica
Gemini	Gemini 2.5 Pro	Pro (pago mensual individual de \$20 USD)	Texto, imágenes, documentos de Google	Integración con Google Workspace, fluidez operativa	Dependiente del ecosistema Google, limitado en contenido técnico	Medio - automatización de contenido, generación en entorno Google
DeepSeek	DeepSeek-V3-0324	Gratis (modelo open source)	Solo texto (estructura técnica)	Precisión técnica, lenguaje estructurado	Curva de aprendizaje alta, no apto para tareas operativas	Alto - redacción técnica, lógica formal, lenguaje especializado

Infinity AI	Infinity AI 1.3.7	Gratis (sin versión Pro disponible)	Texto simple	Agilidad operativa, simplicidad de uso	Capacidad limitada, sin personalización ni análisis avanzado	Bajo - solo tareas simples y operativas
-------------	----------------------	--	--------------	---	--	---

Nota. Comparativa de cinco herramientas GenAI, destacando acceso, multimodalidad, pros y contras, para entender las diferencias entre las herramientas.

CAPÍTULO III – EVALUACIÓN DEL CASO DE ESTUDIO

3.1 INTRODUCCIÓN AL CASO DE ESTUDIO

3.1.1 Descripción del proyecto

El proyecto de caso de uso utilizado en este estudio proviene de una empresa especializada en el desarrollo de soluciones avanzadas basadas en inteligencia artificial, con un enfoque particular en tecnologías de visión por computadora. Su giro de negocio se centra en mejorar la seguridad digital y física mediante herramientas que permiten la autenticación continua de usuarios sin necesidad de contacto, promoviendo experiencias más seguras, privadas y eficientes. Su producto principal es una solución de verificación biométrica en tiempo real que utiliza el rostro del usuario para validar su identidad mientras trabaja, bloqueando automáticamente el dispositivo cuando detecta que el usuario se ha retirado. Todo este proceso ocurre de forma local en el dispositivo, sin enviar datos biométricos a la nube, garantizando así altos estándares de privacidad.

Además de la autenticación continua, su tecnología incluye mecanismos inteligentes para proteger la información visual en pantalla ante observadores no autorizados (como en casos de shoulder surfing) y emplea métodos avanzados de detección de vida para prevenir accesos fraudulentos mediante imágenes o grabaciones. Esta solución está diseñada para integrarse de manera fluida en entornos laborales sensibles, tales como instituciones financieras, gubernamentales o sanitarias, donde la seguridad y la protección de datos son prioridades críticas. Gracias a su enfoque en la privacidad, la eficiencia operativa y la facilidad de uso, esta empresa se posiciona como un referente en la implementación de inteligencia artificial generativa aplicada a la seguridad en tiempo real.

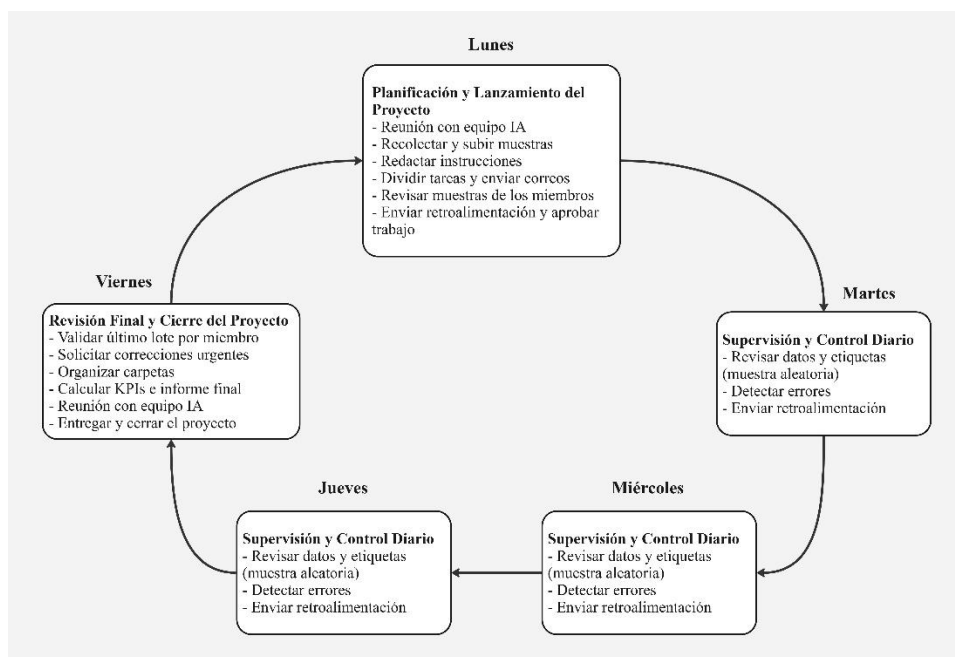
Como parte de sus principios operativos, la empresa prioriza la eficiencia y la diversidad en el desarrollo de sus modelos. Por esta razón, cuenta con un equipo interno conformado por personas con distintas características faciales. Este enfoque garantiza que los datos recolectados sean diversos, representativos y adaptables a las necesidades específicas de los modelos de IA. Asimismo, dado que uno de los pilares fundamentales de la empresa es la protección de la privacidad, todo el proceso de recolección y etiquetado de datos cumple estrictamente con criterios de confidencialidad. De esta forma, se reduce significativamente la dependencia de proveedores externos de datos, lo que no solo disminuye los costos, sino que también permite una mayor personalización, control de calidad y volumen de datos mes a mes.

El caso de estudio sobre el cual se basa esta investigación corresponde a un proyecto semanal real de recolección y etiquetado de imágenes, desarrollado en el contexto de un proceso de alimentación de datos para modelos de inteligencia artificial, específicamente en el área de visión por computadora. Este proyecto forma parte de una etapa crítica conocida como data pipeline, esencial para entrenar modelos con información precisa y de calidad.

Cada semana, un equipo de nueve recolectores y etiquetadores, junto con un supervisor, trabaja con el objetivo de recolectar y etiquetar exactamente 2,500 imágenes, cumpliendo estrictamente con criterios de calidad, precisión en las anotaciones y tiempos de entrega. El ciclo del proyecto se repite de manera constante, iniciando cada lunes y concluyendo cada viernes, permitiendo una planificación operativa bien definida y un flujo de trabajo altamente estandarizado. Este ciclo semanal es el mismo para todos los proyectos del mismo tipo.

Figura 1

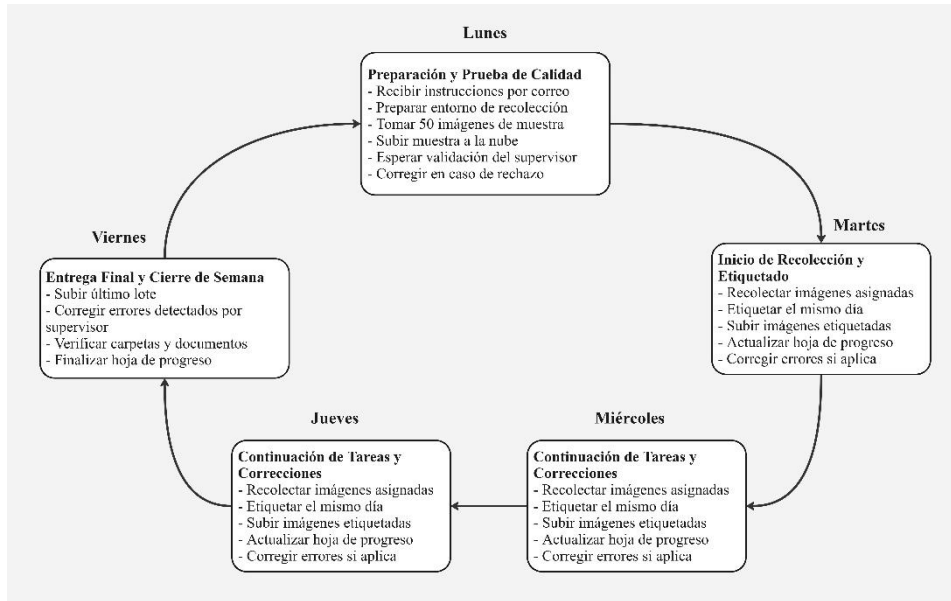
Flujo semanal del supervisor del proyecto



Nota. Muestra el ciclo semanal de supervisión: planificación el lunes, control diario de martes a jueves y cierre el viernes.

Figura 2

Flujo semanal del equipo de recolección y etiquetado



Nota. Ciclo semanal del equipo: prueba y preparación el lunes, recolección y etiquetado de martes a jueves, y cierre de semana el viernes.

El proyecto está estructurado en fases semanales que incluyen: planificación, pruebas de muestra, ejecución diaria, revisión de calidad, corrección de errores y entrega final. Esta organización permite una ejecución consistente y facilita el monitoreo del rendimiento del equipo a través de métricas cuantificables.

El supervisor desempeña un rol clave desde la fase de arranque del proyecto. Cada lunes, mantiene una reunión con el equipo de IA para recibir los requerimientos del proyecto. Posteriormente, redacta las instrucciones de recolección y etiquetado, captura un conjunto de 50 imágenes de muestra como referencia y prepara los materiales informativos (documentación y videos explicativos). Una vez completada esta fase, se asignan tareas equitativas a los nueve miembros del equipo, y se envían los correos de inicio con todos los recursos necesarios. Además, el supervisor realiza revisiones diarias aleatorias de muestras de trabajo (20 imágenes por miembro), proporciona retroalimentación, gestiona correcciones en tiempo real y prepara informes semanales con indicadores clave de desempeño (KPIs).

Por su parte, los miembros del equipo deben trabajar en función de metas diarias, más que de un horario rígido. Aunque están contratados bajo un régimen de media jornada (cuatro horas diarias), la metodología se basa en un modelo de trabajo por objetivos. En la práctica, las

tareas suelen completarse en menos de las 4 horas asignadas. Cada lunes, deben capturar 50 imágenes de muestra como prueba inicial. Estas no forman parte del conteo oficial de las 2,500 imágenes semanales y deben ser aprobadas por el supervisor antes de proceder. De martes a jueves, cada miembro debe recolectar y etiquetar una cantidad fija de imágenes diariamente — ambas tareas deben realizarse el mismo día— y subir los resultados a la carpeta compartida en la nube. Además, deben completar un reporte diario donde registran su progreso, el tiempo estimado dedicado, y cualquier incidencia detectada. Los viernes se realiza la entrega final y, en caso de errores, deben ser corregidos y reemplazados el mismo día. En situaciones críticas, se pueden convocar reuniones urgentes de cierre.

Los datos recolectados son almacenados en Google Drive, plataforma que permite el acceso compartido entre el supervisor y el equipo de IA encargado de entrenar los modelos. También se utilizan herramientas como Google Sheets para el seguimiento diario, correo electrónico y Slack para la comunicación continua.

Este proyecto no solo requiere disciplina y eficiencia operativa, sino también una coordinación activa y fluida. La calidad del etiquetado debe mantenerse alta en todo momento, ya que los errores podrían comprometer el rendimiento del modelo de IA entrenado con estos datos. Por ello, la revisión diaria, la retroalimentación inmediata y la corrección oportuna de errores forman parte integral del proceso.

Los stakeholders involucrados incluyen al supervisor (que lidera el proyecto y consolida el informe de calidad), los nueve miembros del equipo de etiquetado (que cumplen metas diarias según las instrucciones proporcionadas), y el equipo de IA que valida el dataset final para entrenamiento. Todos trabajan sobre una infraestructura técnica básica pero suficiente: cámaras o webcams, computadoras personales, herramientas de anotación, y plataformas colaborativas como Google Drive y Slack.

Riesgos operativos considerados incluyen: errores de etiquetado, incumplimiento de plazos por causas técnicas o personales, fallas en la comunicación, problemas técnicos con los dispositivos o la nube, y la posible pérdida de datos por eliminación accidental o errores de sincronización. Para mitigar estos riesgos, se ha establecido un sistema de control de calidad basado en revisiones aleatorias diarias, corrección inmediata de errores, y reportes individuales por parte de los miembros del equipo.

Finalmente, debido a su carácter repetitivo, colaborativo y basado en resultados medibles, este proyecto constituye un entorno ideal para evaluar el impacto de herramientas de Inteligencia

Artificial Generativa (GenAI). La integración de estas herramientas se realizó en cinco ciclos semanales consecutivos, aplicando una herramienta diferente en cada semana. Cada herramienta analizó la documentación y datos del proyecto y proporcionó recomendaciones que fueron implementadas en tiempo real. Esto permitió observar el efecto directo de cada solución GenAI sobre variables como la organización del trabajo, la eficiencia del equipo y la calidad de los entregables.

Gracias a su diseño metódico y su ejecución en un contexto real, este proyecto proporcionó un marco óptimo para analizar de forma comparativa la utilidad práctica de distintas herramientas de GenAI en la gestión operativa de microproyectos.

3.2 DISEÑO DEL ESTUDIO COMPARATIVO

3.2.1 Metodología de comparación y criterios de evaluación

La evaluación comparativa de las herramientas de Inteligencia Artificial Generativa (GenAI) se diseñó para medir de forma objetiva su impacto en la eficiencia y calidad de gestión de un proyecto real. Esta comparación se fundamenta en un enfoque cuantitativo y estructurado, basado en la recolección sistemática de datos de desempeño, tanto a nivel de equipo como de supervisión, durante cinco semanas consecutivas.

Enfoque general de comparación

Durante cada semana del estudio, se aplicó una herramienta de GenAI diferente al mismo entorno operativo, utilizando documentación real del proyecto para generar recomendaciones de mejora. Dichas recomendaciones fueron implementadas en tiempo real, y su efecto fue monitoreado mediante indicadores clave de desempeño (KPIs) previamente definidos. Se tomaron en cuenta dos niveles de análisis:

1. Rendimiento del equipo de recolección y etiquetado (9 miembros).
2. Desempeño de la supervisión y eficiencia en la gestión del proyecto.

Ambos niveles fueron documentados utilizando hojas de cálculo diarias, reportes semanales, y formularios internos que capturaron tanto variables cuantitativas como apreciaciones cualitativas sobre el flujo de trabajo y la organización.

Instrumentos y métodos de recolección de datos

Para garantizar la fiabilidad de los datos comparativos, se utilizaron los siguientes instrumentos y mecanismos de seguimiento:

- Registro diario del equipo en Google Sheets: Cada miembro registró el número de imágenes recolectadas, el tiempo empleado y cualquier incidencia.
- Informe semanal del supervisor: Incluyó métricas globales, errores detectados, tiempos de revisión, y estadísticas de cumplimiento.
- Control de calidad diario: Revisión aleatoria de 20 imágenes por miembro para garantizar la precisión y coherencia en el etiquetado.
- Al final de cada semana, se aplicaron dos encuestas diferenciadas: una para los miembros del equipo, centrada en evaluar la claridad de las instrucciones, la organización del trabajo y la utilidad práctica de las recomendaciones generadas; y otra para el supervisor, enfocada en valorar la calidad de dichas sugerencias, su aplicabilidad en la gestión del proyecto y su impacto en el desempeño del equipo.

Justificación de la comparabilidad entre semanas

Con el fin de garantizar una evaluación objetiva del impacto de cada herramienta de Inteligencia Artificial Generativa (GenAI), se estableció un marco de trabajo estandarizado para todas las semanas del estudio. Cada semana contó con las mismas condiciones operativas: un equipo de 9 miembros con metas diarias iguales, un supervisor con funciones constantes, el mismo tipo de tarea (recolección y etiquetado de imágenes). La única variable cambiante fue la herramienta de GenAI aplicada. Esta homogeneidad metodológica permite que los resultados obtenidos sean comparables entre sí, eliminando factores de confusión y asegurando que las diferencias observadas en desempeño, eficiencia o precisión puedan atribuirse directamente a la herramienta evaluada en cada ciclo.

Adicionalmente, cada herramienta fue evaluada por el supervisor no solo por su impacto operativo, sino también por la calidad técnica y comunicativa de sus respuestas. Para ello, se registraron aspectos como el tiempo que cada herramienta tomó en analizar la información proporcionada y generar su respuesta, la claridad en la organización del contenido (uso de encabezados, secciones, listas), la adecuación del tono al contexto del proyecto, así como la inclusión de cálculos, proyecciones justificadas y acciones prácticas alineadas con los objetivos reales. Estas observaciones cualitativas fueron complementadas con una valoración estructurada realizada por el supervisor, basada en cinco criterios: relevancia, precisión de los datos, ejecutabilidad, completitud y claridad/estructura, cada uno calificado en una escala de uno a cinco. Esta evaluación integral permitió no solo medir el efecto observable de cada

implementación, sino también analizar el valor técnico y estratégico de las recomendaciones como apoyo real en la toma de decisiones.

Criterios de selección de KPIs

Los indicadores clave de desempeño (KPIs) utilizados en este estudio fueron seleccionados con base en su relevancia directa para evaluar el rendimiento real en un entorno de trabajo colaborativo. Se priorizaron métricas que pudieran capturar de forma cuantitativa aspectos esenciales de la productividad (como la velocidad de procesamiento), la calidad (como la tasa de error y precisión), la eficiencia en la gestión (como el tiempo promedio por muestra) y la inversión de recursos (como el costo por imagen). Estos KPIs fueron definidos desde la fase inicial del estudio y aplicados de manera uniforme en todas las semanas, garantizando la posibilidad de realizar comparaciones consistentes y objetivas.

Indicadores de desempeño del supervisor del proyecto

Como parte del enfoque metodológico, se evaluaron las actividades del supervisor a lo largo de la semana, considerando su rol central en la organización, planificación, revisión y control de calidad del proyecto. Para ello, se definieron indicadores específicos que permiten medir el tiempo invertido en tareas clave, así como la eficiencia en la gestión general. La siguiente tabla muestra los principales KPIs semanales relacionados con la labor del supervisor. Estos datos fueron calculados a partir del desglose de actividades realizadas entre lunes y viernes, según los registros operativos del proyecto.

Tabla 3

Indicadores de desempeño semanal del supervisor del proyecto

Indicador	Fórmula / Descripción	Valor semanal	Unidad
Tiempo total de supervisión	Suma de todas las actividades realizadas por el supervisor de lunes a viernes	1390	Minutos
Tiempo total de planificación	Tiempo de planificación del proyecto semanal	200	Minutos
Eficiencia de planificación	$(\text{Tiempo de planificación} / \text{Tiempo total de supervisión}) \times 100$	14.39	Porcentaje
Número de muestras revisadas	Suma de muestras revisadas entre martes y jueves	1440	Número de archivos
Tiempo dedicado a revisión diaria	Suma del tiempo destinado a revisar las muestras por día entre martes y jueves	630	Minutos
Tiempo promedio por muestra	Tiempo total de revisión / número de muestras revisadas	0.44	Minutos/muestra

Nota. Resume los principales KPIs del supervisor, detallando tiempos, eficiencia y volumen de muestras para evaluar su desempeño semanal.

Estos indicadores permiten visualizar de forma clara el nivel de carga operativa y el grado de organización del trabajo del supervisor. A partir de estos valores, se podrá analizar cómo las herramientas de GenAI aplicadas en semanas posteriores afectan la distribución del tiempo, la eficiencia en la planificación y la calidad del control ejercido sobre el equipo.

Rendimiento semanal del equipo: indicadores clave (KPIs)

En paralelo a la evaluación del supervisor, se realizó un seguimiento cuantitativo del desempeño del equipo de recolección y etiquetado, compuesto por 9 miembros. Los indicadores clave de desempeño (KPIs) fueron calculados semanalmente en función de métricas operativas como la cantidad de imágenes procesadas, los tiempos invertidos, la precisión en la ejecución de tareas y el cumplimiento de plazos.

Tabla 4

Indicadores operativos semanales del equipo de recolección y etiquetado

Tipo	Indicador	Fórmula / Descripción	Valor semanal	Unidad
Recolección	Total de imágenes recolectadas	Suma de todas las imágenes recolectadas por el equipo	2,500	Número de imágenes
	Tasa de cumplimiento	Porcentaje de miembros que completaron su meta diaria	100%	Porcentaje
	Tiempo total en recolección	Tiempo total invertido en recolección por el equipo	1914	Minutos
	Velocidad de recolección	Promedio de imágenes recolectadas por minuto	1.35	Número de imágenes/min
	Errores en recolección	Total, de errores detectados en la recolección	35	Número de errores
	Tasa de error en recolección	$(\text{Errores} / \text{Total recolectado}) \times 100$	1.40%	Porcentaje
	Precisión de la recolección	$1 - (\text{Errores} / \text{Total recolectado}) \times 100$	98.60%	Porcentaje
Etiquetado	Total de imágenes etiquetadas	Suma de imágenes etiquetadas por el equipo	2,500	Número de imágenes
	Tasa de cumplimiento	Porcentaje de miembros que completaron su meta diaria	100%	Porcentaje
	Tiempo total en etiquetado	Tiempo total invertido en etiquetado por el equipo	3273	Minutos
	Velocidad de etiquetado	Promedio de imágenes etiquetadas por minuto	0.77	Número de imágenes/min
	Errores en etiquetado	Total, de errores detectados en las anotaciones	47	Número de errores
	Tasa de error en etiquetado	$(\text{Errores} / \text{Total etiquetado}) \times 100$	1.88%	Porcentaje
	Precisión del etiquetado	$1 - (\text{Errores} / \text{Total etiquetado}) \times 100$	98.12%	Porcentaje
General	Total de miembros	Número total de personas que integran el equipo.	9	Número de personas

Productividad general	(Total recolectado + etiquetado) / Total de horas trabajadas	52.1	Número de imágenes/hora
Tiempo total invertido	Suma de horas trabajadas por el equipo durante la semana	97	Horas
Tiempo total de subida	Suma del tiempo empleado en subir archivos	633	Minutos
Costo por Imagen	Costo promedio por imagen.	0.22	Dólares
Tiempo promedio de subida por archivo	(Tiempo de subida / total de archivos)	0.13	Minutos/imagen

Nota. Indicadores clave de recolección y etiquetado: volúmenes, tiempos, velocidades, errores, precisión, productividad y costo.

Es importante señalar que los datos presentados en esta tabla corresponden al periodo de martes a viernes, ya que los días lunes están destinados a tareas preparatorias. Estas incluyen la lectura y comprensión de las instrucciones, la organización del entorno de trabajo, la recopilación de datos de muestra (50 imágenes por miembro) y el envío de dichos archivos al supervisor para su revisión y aprobación. El tiempo total invertido por los 9 miembros en estas tareas fue de 715 minutos, lo que equivale aproximadamente a 11.9 horas de trabajo conjunto.

La información aquí registrada representa el rendimiento consolidado del equipo durante los días de ejecución operativa, y servirá como línea base para comparar el impacto de cada herramienta de Inteligencia Artificial Generativa aplicada en semanas posteriores.

Estimación aproximada del presupuesto mensual del proyecto

Para complementar el análisis del rendimiento del equipo y la supervisión, se presenta a continuación una estimación del presupuesto mensual requerido para llevar a cabo este tipo de proyecto en condiciones similares. El costo mensual estimado contempla la participación de un supervisor, nueve miembros del equipo de recolección y etiquetado, y el uso de almacenamiento en la nube para la gestión de archivos.

Tabla 5

Estimación de costos mensuales del proyecto

Concepto	Costo mensual (USD)
Supervisor	\$800.00
9 recolectores/etiquetadores	\$3,600.00
Almacenamiento en la nube	\$18.40
Herramientas y software	\$0.00
Total mensual estimado	\$4,418.40

Nota. Desglose claro de los costos mensuales por rol y recursos.

En conjunto, la metodología aplicada en este estudio permite establecer una base sólida para la evaluación comparativa entre distintas herramientas de Inteligencia Artificial Generativa. A través de la estandarización de tareas semanales, la definición clara de KPIs, la documentación rigurosa de actividades tanto del equipo como del supervisor, y el control de costos operativos, se ha construido un entorno controlado que facilita el análisis objetivo del impacto de cada herramienta en la eficiencia, la calidad del trabajo y la gestión del tiempo y los recursos. Los datos obtenidos durante esta primera etapa constituyen un punto de referencia fundamental para interpretar los cambios que puedan producirse en semanas posteriores, cuando se integren distintas soluciones de GenAI al flujo de trabajo real del proyecto.

3.3 PROTOCOLO DE IMPLEMENTACIÓN Y ESTANDARIZACIÓN DE INSTRUCCIONES

3.3.1 Esquema de implementación

Con el objetivo de evaluar el impacto real de herramientas de Inteligencia Artificial Generativa (GenAI) en la gestión de proyectos, se diseñó un proceso de implementación gradual y controlado, directamente aplicado al caso de estudio presentado anteriormente. Esta fase consistió en aplicar, de manera individual, cinco herramientas GenAI diferentes a lo largo de cinco semanas consecutivas del mismo proyecto, manteniendo una estructura operativa constante para asegurar la comparabilidad de los resultados.

Durante cada semana, se seleccionó una herramienta distinta de GenAI y se le proporcionó un conjunto completo de información sobre el proyecto. Esto incluyó una descripción detallada del caso de uso, el flujo de trabajo operativo del equipo, las restricciones de tiempo, los indicadores clave de desempeño (KPIs) obtenidos de manera tradicional, registrados durante la ejecución previa al uso de cualquier herramienta GenAI. Estos datos fueron utilizados como insumo base en cada semana para que las herramientas pudieran generar recomendaciones personalizadas. A partir de esta información, se solicitó a cada herramienta que ofreciera recomendaciones específicas, prácticas y directamente aplicables orientadas a mejorar cuatro áreas fundamentales del proyecto:

1. La optimización del proceso,
2. El uso eficiente de los recursos del equipo,
3. El incremento de la productividad colectiva,
4. Y la eficiencia en el rol del supervisor.

Cada herramienta recibió la misma instrucción base, junto con los archivos de datos correspondientes, y fue evaluada principalmente en función de la calidad, claridad, aplicabilidad y valor estratégico de sus recomendaciones. Para mantener la equidad metodológica, se priorizó la evaluación de la primera respuesta generada por cada herramienta, simulando una interacción realista y directa. No se utilizaron herramientas externas ni se introdujeron cambios estructurales en la operación semanal. Además, todas las herramientas fueron instruidas para generar recomendaciones únicamente dentro del marco de recursos ya disponibles, sin asumir aumentos de personal, cambios tecnológicos ni extensiones horarias. De este modo, se garantizó que todas las mejoras propuestas pudieran ser implementadas de forma inmediata dentro del flujo de trabajo existente.

Este enfoque permite observar cómo cada herramienta de GenAI interpreta la información, identifica oportunidades de mejora, y propone estrategias viables en un entorno operativo real. En las siguientes secciones se presenta el análisis de resultados semanales obtenidos con cada herramienta aplicada, permitiendo evaluar de forma comparativa su utilidad, precisión y potencial de integración en la gestión de proyectos.

3.3.2 Instrucciones proporcionadas a las herramientas GenAI

Durante la implementación semanal, cada herramienta de GenAI recibió un conjunto estandarizado de instrucciones con el fin de garantizar una evaluación equitativa, objetiva y replicable. Este conjunto se estructuró en forma de un prompt detallado que simulaba una situación real: un supervisor de proyecto solicitando asesoría directa a una herramienta GenAI para planificar la siguiente semana operativa del equipo.

El objetivo fue asegurar que todas las herramientas evaluadas trabajaran bajo las mismas condiciones y con los mismos insumos, permitiendo así comparar la calidad, aplicabilidad y valor estratégico de las respuestas obtenidas.

Contenido del prompt

El prompt entregado a cada herramienta incluyó los siguientes elementos:

- Descripción del Proyecto: se explicaba que el estudio se basa en un proyecto de recolección y etiquetado de imágenes ejecutado por un equipo de 9 miembros, supervisado por un coordinador general, con una meta de 2,500 imágenes por semana. Se detalló la estructura operativa semanal (planificación los lunes,

ejecución de martes a viernes), el entorno de trabajo remoto y la distribución de tareas entre los miembros y el supervisor.

- Indicadores Clave de Desempeño (KPIs): se adjuntaron los valores reales correspondientes a la semana inmediatamente anterior a la implementación de las herramientas GenAI. Estos incluyen la productividad por hora, las tasas de error, los tiempos invertidos en cada tarea y la carga de trabajo del supervisor.
- Diagramas de Flujo: se incluyeron esquemas visuales del flujo operativo tanto del equipo como del supervisor para contextualizar las interacciones semanales.
- Condiciones Operativas Obligatorias:
 - Todos los miembros deben realizar tanto la recolección como el etiquetado de sus imágenes (no es posible dividir funciones).
 - Cada miembro dispone de hasta cuatro horas diarias, sin tiempos fijos por tarea.
 - No se permite proponer nuevas herramientas, software, personal adicional o cambios tecnológicos.
 - Las recomendaciones deben adaptarse exclusivamente al flujo de trabajo y recursos ya disponibles.
- Formato y Alcance Esperado de la Respuesta:
 - Las recomendaciones debían organizarse en torno a cuatro áreas clave:
 1. Optimización del proceso: en este apartado, la IA debe identificar pasos que se repitan o que estén fragmentados de forma innecesaria dentro del ciclo semanal. La respuesta debe proponer cómo agrupar tareas que actualmente se ejecutan por separado o reordenar la secuencia de actividades para eliminar solapamientos. Cada sugerencia debe ir acompañada de una descripción clara de la lógica detrás del ahorro de tiempo y de los beneficios operativos que se obtendrían al implementar esos ajustes.
 2. Uso eficiente de los recursos del equipo: aquí se espera que la IA analice la capacidad de cada colaborador en función de sus tiempos de trabajo habituales y sugiera una nueva distribución de tareas que equilibre la carga según su rendimiento. También debe proponer actividades complementarias que mejoren la calidad del trabajo sin incrementar la supervisión directa. Las recomendaciones deben fundamentarse en el

historial de desempeño y en la necesidad de maximizar el aprovechamiento del tiempo disponible.

3. Mejora de la productividad colectiva: este bloque se enfoca en detectar cuellos de botella, tales como procesos que ralentizan el flujo general, y en detallar acciones concretas para agilizar esos puntos críticos. Deberá incluirse una estimación del impacto cualitativo de cada acción sobre la velocidad de trabajo y la consistencia en los resultados, describiendo cómo cambiarían las dinámicas diarias al adoptar las propuestas.

4. Eficiencia del supervisor: la IA debe evaluar la distribución actual del trabajo del supervisor —planificación, revisión y comunicación— y recomendar herramientas y técnicas que aceleren las tareas rutinarias. Cada sugerencia debe explicar cómo permite liberar tiempo de actividades repetitivas y de qué manera ese tiempo liberado podría redirigirse a tareas de mayor valor estratégico, sin necesidad de introducir nuevas plataformas o roles.

- Incluir cálculos reales, proyecciones numéricas y propuestas específicas de implementación.
- Proporcionar un plan de acción semanal, con priorización de tres acciones clave y justificación basada en los datos entregados.

Se enfatizó que no se requerían consejos generales ni buenas prácticas genéricas: todas las recomendaciones debían derivarse de los datos adjuntos y orientarse a resultados cuantificables, listos para aplicar.

Archivos adjuntos para el análisis

Junto con el prompt, cada herramienta recibió los siguientes documentos:

- **Managment Performance.xlsx**: métricas reales del supervisor, incluyendo tiempos de planificación, revisión, retroalimentación y tareas administrativas.
- **Team Performance.xlsx**: rendimiento individual del equipo por jornada, incluyendo imágenes procesadas por hora, tasa de errores y tiempo efectivo trabajado.
- **Weekly Flow of the Collection and Labeling Equipment.jpg**: diagrama del flujo operativo del equipo.

- Weekly flow of the project supervisor.jpg: visualización del ciclo semanal del supervisor.

Criterio metodológico

Todas las herramientas recibieron exactamente el mismo conjunto de información, y se evaluó únicamente su primera respuesta generada, sin ajustes ni retroalimentación posterior. Este enfoque buscó simular una interacción realista y directa entre un profesional y una herramienta GenAI, donde se requiere una solución clara y aplicable en un solo intento.

3.4 IMPLEMENTACIÓN Y EVALUACIÓN DE HERRAMIENTAS DE GENAI

En esta sección se presentan los resultados obtenidos tras aplicar, de forma individual, cada una de las cinco herramientas de Inteligencia Artificial Generativa seleccionadas.

3.4.1 Resultados de la implementación con ChatGPT

Durante la semana de implementación (del 28 de abril al 2 de mayo de 2025), se empleó ChatGPT Pro con el modelo GPT-4o, que procesó sin dificultades archivos Excel, imágenes JPG y texto estructurado. El análisis requirió solo 12 segundos y la generación de recomendaciones, 9 segundos. La respuesta se presentó en secciones claramente delimitadas —con encabezados y subtítulos por área de mejora—, e incluyó cálculos precisos y proyecciones ajustadas a los datos reales, cada uno acompañado de su respectiva justificación. Cada bloque combinó un diagnóstico de la situación actual con acciones concretas, expresadas de forma directa y práctica, y añadió la opción de entregar los resultados en PDF o presentación para facilitar su implementación y seguimiento. El lenguaje, técnico pero accesible, permitió una comprensión fluida incluso sin formación especializada. En conjunto, ChatGPT demostró gran agilidad en el procesamiento, rigor cuantitativo y un estilo humanizado, orientando con claridad las mejoras inmediatas a incorporar en el flujo de trabajo.

Resumen de recomendaciones

1. Optimización del proceso:

ChatGPT identificó tareas redundantes dentro del flujo diario, como la actualización constante de hojas de progreso. Propuso concentrar la carga de imágenes al final del día. También recomendó registrar el avance solo lunes y viernes, eliminando entradas innecesarias los martes a jueves.

2. Uso eficiente de los recursos del equipo:

Con una capacidad semanal de 180 horas disponibles (9 miembros × 20 horas), el uso efectivo fue solo del 53.9% (≈97 horas reales). Se recomendó aumentar las imágenes asignadas por persona de 278 a 300 (un +7.9%) para ajustar la carga al ritmo actual sin sobrecarga adicional.

3. Mejora de la productividad del equipo:

La fase de etiquetado (0.77 imágenes/min) fue detectada como el cuello de botella frente a la recolección (1.35 imágenes/min). ChatGPT sugirió agrupar las actividades: recolección martes y miércoles, etiquetado jueves y viernes, con el fin de aumentar la velocidad de etiquetado a 0.9 imágenes/min. También propuso centralizar la retroalimentación en sesiones grupales con ejemplos comunes.

4. Eficiencia del supervisor:

Se registraron 1,390 minutos semanales de supervisión, de los cuales solo 200 minutos correspondían a planificación, representando una eficiencia del 14.4%. Se propuso concentrar toda la planificación en un bloque de 2 horas los lunes, y revisar 20 muestras en vez de 40 por miembro con el objetivo de aumentar la eficiencia, reduciendo microajustes y liberando tiempo operativo.

5. Plan de acción sugerido por ChatGPT:

- Lunes: Planificación completa del proyecto y revisión de muestras iniciales.
- Martes y miércoles: Solo recolección de imágenes.
- Jueves: Solo etiquetado.
- Viernes: Cierre de tareas, retroalimentación grupal y carga final.

El supervisor valoró la respuesta de ChatGPT con un 4/5 en relevancia, 3/5 en precisión de datos, 5/5 en ejecutabilidad, completitud y en claridad y estructura. Implementó todas las recomendaciones: concentró la subida de imágenes al cierre de la jornada, unificó el registro de progreso solo los lunes y viernes, ajustó la meta semanal de 278 a 300 imágenes por colaborador, reorganizó las fases de trabajo (recolección martes–miércoles y etiquetado jueves–viernes), consolidó la retroalimentación en sesiones grupales, centralizó toda la planificación en una sesión intensiva el lunes y evitó ajustes menores durante la semana.

KPIs obtenidos vs. semana base:

Tabla 6

Comparación de indicadores de desempeño del supervisor: semana base vs. uso de ChatGPT

Indicador	Valor semanal	Valor con ChatGPT	Variación absoluta	Unidad
Tiempo total de supervisión	1390	980	-410	Minutos
Tiempo total de planificación	200	200	0	Minutos
Eficiencia de planificación	14.39%	20.41%	6.02%	Porcentaje
Número de muestras revisadas	1440	720	-720	Número de archivos
Tiempo dedicado a revisión diaria	630	490	-140	Minutos
Tiempo promedio por muestra	0.44	0.68	0.24	Minutos/muestra

Nota. Comparación de indicadores clave antes y después de integrar ChatGPT, mostrando el impacto en tiempos, eficiencia y volumen de trabajo.

Tabla 7

Comparación de indicadores de desempeño del equipo de recolección y etiquetado: semana base vs. uso de ChatGPT

Tipo	Indicador	Valor original	Valor con ChatGPT	Variación absoluta	Unidad
Recolección	Total de imágenes recolectadas	2500	2,700	200	Número de imágenes
	Tasa de cumplimiento	100%	100%	0%	Porcentaje
	Tiempo total en recolección	1914	2275	361	Minutos
	Velocidad de recolección	1.35	1.20	-0.15	Imágenes/min
	Errores en recolección	35	41	6	Número de errores
	Tasa de error en recolección	1.40%	1.52%	0.12%	Porcentaje
	Precisión de la recolección	98.60%	98.48%	-0.12%	Porcentaje
Etiquetado	Total de imágenes etiquetadas	2500	2,700	200	Número de imágenes
	Tasa de cumplimiento	100%	100%	0%	Porcentaje
	Tiempo total en etiquetado	3273	3138	-135	Minutos
	Velocidad de etiquetado	0.77	0.86	0.09	Imágenes/min
	Errores en etiquetado	47	25	-22	Número de errores
	Tasa de error en etiquetado	1.88%	0.93%	-0.95%	Porcentaje
	Precisión del etiquetado	98.12%	99.07%	0.95%	Porcentaje
General	Total de miembros	9	9	0	Número de personas
	Productividad general	51.55	56.47	4.92	Imágenes/hora
	Tiempo total invertido	97.00	95.63	-1.37	Horas
	Tiempo total de subida	633	325	-308	Minutos
	Costo por Imagen	0.22	0.20	-0.02	Dólares
	Tiempo promedio de subida por archivo	0.13	0.06	-0.07	Minutos/imagen

Nota. Comparativa de desempeño antes y después de usar ChatGPT en recolección y etiquetado, resaltando cambios en volúmenes, velocidades, precisión y costos.

Durante la semana de implementación con ChatGPT, el tiempo total de supervisión se redujo significativamente de 1,390 a 980 minutos, destacando una disminución de 140 minutos en las tareas de revisión diaria. Aunque el número de muestras revisadas se redujo a la mitad (de 1,440 a 720), el tiempo promedio por muestra aumentó ligeramente de 0.44 a 0.68 minutos. La eficiencia de planificación semanal también mejoró, pasando del 14.39% al 20.41%, cumpliendo con el objetivo de concentrar la planificación en un solo bloque y reducir microajustes durante la semana. En cuanto al equipo, se observó un aumento en la productividad general, pasando de 51.55 a 56.47 imágenes por hora. La velocidad de etiquetado mejoró de 0.77 a 0.86 imágenes por minuto y la tasa de error bajó del 1.88% al 0.93%, con una reducción de 135 minutos en el tiempo total dedicado a esta tarea. El tiempo de subida se redujo más de la mitad (de 633 a 325 minutos), disminuyendo también el costo por imagen. Aunque la recolección de imágenes aumentó de 2,500 a 2,700 unidades, la velocidad en esta fase se redujo ligeramente y los errores aumentaron levemente, reflejando la necesidad de revisar el ritmo frente a la calidad.

Percepción del uso

Desde la perspectiva del supervisor, las respuestas fueron mayoritariamente positivas. Se indicó un claro alineamiento entre las recomendaciones generadas por ChatGPT y las condiciones reales del proyecto, así como con sus objetivos específicos. En particular, se valoró altamente la capacidad de la herramienta para ofrecer recomendaciones claras, coherentes y directamente aplicables. También se reconoció que el uso de ChatGPT permitió detectar ineficiencias previamente no identificadas y que sus sugerencias aportaron valor a la planificación y a la toma de decisiones. Aunque la mejora en la calidad del trabajo entregado y la fluidez en la supervisión recibieron una valoración más neutral, la herramienta fue percibida como fácil de usar y con una interfaz intuitiva. En cuanto al equipo, las percepciones también fueron en general favorables. La mayoría de los miembros consideró que las tareas asignadas estaban claras desde el inicio y que la planificación diaria contribuyó al cumplimiento de sus metas. Se evidenció una mejora en la percepción de productividad personal y en la organización del flujo de trabajo en comparación con semanas anteriores. La mayoría también evaluó positivamente la carga de trabajo semanal y la eficiencia general en la gestión del proyecto. No obstante, algunas respuestas reflejaron experiencias mixtas respecto a la retroalimentación recibida, con dos miembros indicando desacuerdo o neutralidad en este aspecto.

3.4.2 Resultados de la implementación con Claude

Durante la segunda semana del estudio (del 5 a 9 de mayo de 2025), se implementó la herramienta Claude en su versión Sonnet 4 Pro. Aunque aceptó todos los formatos proporcionados, no admitió archivos Excel directamente, por lo que fue necesario enviar capturas de pantalla de los datos. El análisis de los archivos tomó únicamente 3 segundos, y la generación de recomendaciones se completó en 22 segundos e incluyó una respuesta estructurada en bloques bien diferenciados por áreas estratégicas. El texto se presentó con un lenguaje técnico-formal, claro y directo, con un enfoque colaborativo y propositivo. Las recomendaciones incluyeron proyecciones cuantificadas, estimaciones de mejora en tiempo y rendimiento, así como cálculos específicos sobre capacidad ociosa, redistribución de carga y eficiencia operativa. Además, propuso un plan de acción gradual, detallando el orden de ejecución diaria, la reasignación diferenciada de tareas según desempeño individual y un nuevo esquema de retroalimentación. La respuesta no solo abordó las debilidades detectadas, sino que priorizó la implementación sin introducir cambios tecnológicos ni estructurales, manteniéndose completamente alineada con los objetivos y condiciones reales del proyecto.

Resumen de recomendaciones

1. Optimización del proceso:

Claude sugirió reorganizar las actividades de lunes y viernes para reducir la carga operativa del supervisor. Reemplazó las 50 muestras individuales de revisión por una combinación de 25 individuales y 25 imágenes grupales. Además, propuso eliminar la revisión carpeta por carpeta el viernes, reemplazándola con una validación por lotes que ahorra 90 minutos adicionales.

2. Uso eficiente de los recursos del equipo:

Detectó una subutilización significativa de la capacidad operativa (solo el 60% del tiempo disponible estaba siendo usado). Claude redistribuyó las asignaciones según el rendimiento individual: los miembros más eficientes recibirían 320 imágenes semanales (+15%), los de rendimiento medio 290 (+4%), y los de menor desempeño, una carga reducida de 250. Esta personalización permite aprovechar 72 horas semanales ociosas sin comprometer la calidad.

3. Mejora de la productividad del equipo:

Propuso una estructura horaria diaria dividida en bloques concentrados de 90 minutos para recolección y etiquetado, separados por pausas y seguido por cargas por lotes. Con este

cambio, estimó un aumento de la velocidad de trabajo de 52.1 a más de 65 imágenes/hora. También planteó un nuevo esquema de retroalimentación: feedback grupal entre martes y jueves, con retroalimentación correctiva individual.

4. Eficiencia del supervisor:

Planteó una reducción del tiempo de supervisión de 1,390 a 1,050 minutos semanales (-24%) mediante tres acciones: revisar muestras por lotes los lunes, reducir de 20 a 15 las muestras aleatorias revisadas entre martes y jueves, y dedicar los viernes al análisis estratégico con 60 minutos adicionales para la planificación del siguiente semanal.

5. Plan de acción sugerido por Claude:

- Domingo: Preparar el sistema de subida por lotes y la nueva distribución de tareas.
- Lunes: Lanzamiento del nuevo sistema de muestras y horarios optimizados.
- Martes a jueves: Aplicar feedback grupal y revisar solo 15 muestras diarias.
- Viernes: Evaluación del desempeño semanal y planificación siguiente.

El supervisor valoró la respuesta de Claude con un 4/5 en relevancia, 3/5 en precisión de datos, 4/5 en ejecutabilidad y pleno 5/5 en completitud y claridad estructural. Se implementaron casi todas las recomendaciones, destacando como especialmente útiles la redistribución personalizada de tareas, la simplificación de lunes y viernes, y el nuevo enfoque de retroalimentación. Se elogió la inclusión del bloque de preparación del domingo y la estrategia de implementación gradual. Las únicas observaciones fueron la falta de indicaciones específicas por miembro en el proceso de revisión y las dificultades prácticas para implementar una rutina horaria rígida.

KPIs obtenidos vs. semana base:

Tabla 8

Comparación de indicadores de desempeño del supervisor: semana base vs. uso de Claude

Indicador	Valor semanal	Valor con Claude	Variación absoluta	Unidad
Tiempo total de supervisión	1390	1006	-384	Minutos
Tiempo total de planificación	200	200	0	Minutos
Eficiencia de planificación	14.39%	19.88%	5.49%	Porcentaje
Número de muestras revisadas	1440	540	-900	Número de archivos

Tiempo dedicado a revisión diaria	630	561	-69	Minutos
Tiempo promedio por muestra	0.44	1.04	0.60	Minutos/muestra

Nota. Comparación de indicadores clave antes y después de integrar Claude, mostrando el impacto en tiempos, eficiencia y volumen de trabajo.

Tabla 9

Comparación de indicadores de desempeño del equipo de recolección y etiquetado: semana base vs. uso de Claude

Tipo	Indicador	Valor original	Valor con Claude	Variación absoluta	Unidad
Recolección	Total de imágenes recolectadas	2500	2,632	132	Número de imágenes
	Tasa de cumplimiento	100%	100%	0%	Porcentaje
	Tiempo total en recolección	1914	1756	-158	Minutos
	Velocidad de recolección	1.35	1.58	0.23	Imágenes/min
	Errores en recolección	35	36	1	Número de errores
	Tasa de error en recolección	1.40%	1.37%	-0.03%	Porcentaje
	Precisión de la recolección	98.60%	98.63%	0.03%	Porcentaje
Etiquetado	Total de imágenes etiquetadas	2500	2,632	132	Número de imágenes
	Tasa de cumplimiento	100%	100%	0%	Porcentaje
	Tiempo total en etiquetado	3273	3440	167	Minutos
	Velocidad de etiquetado	0.77	0.77	0.00	Imágenes/min
	Errores en etiquetado	47	54	7	Número de errores
	Tasa de error en etiquetado	1.88%	2.05%	0.17%	Porcentaje
	Precisión del etiquetado	98.12%	97.95%	-0.17%	Porcentaje
General	Total de miembros	9	9	0	Número de personas
	Productividad general	51.55	50.98	-0.56	Imágenes/hora
	Tiempo total invertido	97.00	103.25	6.25	Horas
	Tiempo total de subida	633	999	366	Minutos
	Costo por Imagen	0.22	0.21	-0.01	Dólares
	Tiempo promedio de subida por archivo	0.13	0.19	0.06	Minutos/imagen

Nota. Comparativa de desempeño antes y después de usar Claude en recolección y etiquetado, resaltando cambios en volúmenes, velocidades, precisión y costos.

Durante la semana de implementación con Claude, el tiempo total de supervisión se redujo de 1,390 a 1,006 minutos, principalmente gracias a una reorganización de actividades y validaciones por lotes, lo que también permitió disminuir en 69 minutos el tiempo diario dedicado a revisión. Aunque el número de muestras revisadas bajó significativamente (de 1,440 a 540), el tiempo promedio por muestra se incrementó de 0.44 a 1.04 minutos. La eficiencia de

planificación aumentó de 14.39% a 19.88%, alineándose con la propuesta de dedicar más tiempo a actividades estratégicas los lunes.

En cuanto al equipo, la recolección de imágenes mejoró en velocidad, pasando de 1.35 a 1.58 imágenes por minuto, con una leve reducción en el tiempo total y una mejora marginal en la precisión. Sin embargo, la fase de etiquetado no mostró avances en velocidad y presentó un ligero aumento en los errores (de 47 a 54). La productividad general del equipo bajó levemente de 51.55 a 50.98 imágenes por hora, mientras que el tiempo total invertido aumentó en 6.25 horas. El tiempo de subida creció considerablemente (de 633 a 999 minutos), lo que elevó el promedio por archivo, aunque el costo por imagen se mantuvo estable.

Percepción del uso

Desde la perspectiva del supervisor; se reconoció que las recomendaciones de Claude estuvieron alineadas con las condiciones reales del proyecto y aportaron valor en la organización general y la supervisión. No obstante, aspectos como la claridad y aplicabilidad de las recomendaciones, así como su coherencia con los objetivos específicos, recibieron valoraciones más neutrales. Aun así, la herramienta fue percibida como accesible, con una interfaz comprensible y sin requerimientos técnicos elevados. En cuanto al equipo, las percepciones fueron variadas. Si bien varios miembros destacaron una carga de trabajo manejable y expresaron sentirse más productivos, se identificaron áreas de oportunidad. Por ejemplo, algunos participantes indicaron que la planificación diaria no siempre facilitó el cumplimiento de metas y que la retroalimentación no fue completamente útil o clara. Las respuestas también reflejaron neutralidad en la organización del flujo de trabajo y en la comunicación general. En resumen, aunque Claude ofreció aportes relevantes a nivel operativo, la experiencia de usuario mostró un impacto menos uniforme comparado con la semana anterior.

3.4.3 Resultados de la implementación con Gemini

Durante la tercera semana del estudio (del 12 al 16 de mayo de 2025), se utilizó Gemini 2.5 Pro, que aceptó correctamente todos los formatos enviados, incluyendo hojas Excel, diagramas y texto estructurado. El análisis de los archivos tomó solo 16 segundos y la generación de recomendaciones, 19 segundos. La respuesta se entregó en bloques bien organizados por áreas estratégicas, con encabezados claros y un lenguaje técnico-colaborativo, formal y propositivo. Incluyó estimaciones numéricas de impacto, cálculos proyectados, estrategias diferenciadas por perfil de desempeño y sugerencias prácticas para reducir errores y mejorar la eficiencia del equipo y del supervisor. A diferencia de otras herramientas, Gemini

propuso un enfoque iterativo basado en revisión focalizada, delegación estratégica y checklist de autoevaluación, con un fuerte énfasis en la retroalimentación y la gestión preventiva de errores. Su plan se organizó por días y tareas específicas, y evitó la implementación de herramientas externas o cambios estructurales.

Resumen de recomendaciones

1. Optimización del proceso:

Gemini detectó que la revisión de muestras los lunes podía generar cuellos de botella y propuso priorizar su aprobación en bloques de tiempo definidos, aplicando revisiones tipo spot-check, sugirió consolidar la carga de archivos en una única sesión al final de cada jornada para reducir interrupciones operativas. Estas medidas buscaban ahorrar entre 1.5 y 2 horas diarias del equipo en gestión de archivos.

2. Uso eficiente de los recursos del equipo:

Con base en el análisis del archivo Weekly Performance.csv, Gemini propuso estrategias para definir metas personalizadas según el rendimiento individual, con el objetivo de optimizar la capacidad operativa de cada miembro. La fórmula sugerida para calcular la nueva meta diaria fue: Nueva Meta Diaria = (Imágenes/Hora × 4 horas), considerando un promedio de 4 horas laborales por día. Sin embargo, aunque se presentó una tabla de ejemplo y se incluyó la fórmula, los datos no fueron completados ni calculados con valores reales del proyecto. Aun así, la estrategia buscaba exceder la meta semanal de 2,500 imágenes, apuntando a una producción de aproximadamente 2,700 imágenes para generar un margen de seguridad del 5 al 10%.

3. Mejora de la productividad del equipo:

La herramienta propuso reemplazar revisiones aleatorias por una revisión focalizada el martes por la mañana, complementada con reportes diarios de errores comunes y retroalimentación individual dirigida. También introdujo una checklist de autoevaluación (5–7 puntos) obligatoria antes de subir archivos, y una sesión opcional de Q&A semanal. Estas medidas estaban orientadas a reducir la tasa de error entre un 10 y 15%, promoviendo el aprendizaje autónomo y la consistencia del equipo.

4. Eficiencia del supervisor:

Para reducir el tiempo de supervisión, Gemini planteó revisar por lotes según historial de desempeño, aplicar un modelo de revisión basada en riesgo (enfocando el tiempo en miembros

con errores >3%) y delegar la documentación operativa a un miembro del equipo. Esta estrategia buscaba disminuir entre un 15 y 20% del tiempo dedicado a revisión y liberar hasta 2 horas semanales de tareas administrativas.

5. Plan de acción sugerido por Gemini:

- Domingo: Cálculo y asignación de nuevas metas individuales; selección del miembro para documentar instrucciones.
- Lunes: Preparación inicial, envío y revisión rápida de muestras, aprobación y retroalimentación inicial.
- Martes: Revisión focalizada, espacio Q&A opcional, inicio de recolección masiva.
- Martes a viernes: Recolección, uso de checklist diaria, carga única de archivos, feedback individual.
- Viernes: Análisis de KPIs, evaluación de resultados, resumen para planificación siguiente semana.

El supervisor valoró la respuesta de Gemini con un 3/5 en relevancia, 2/5 en precisión de datos, 3/5 en ejecutabilidad, 5/5 en completitud y 4/5 en claridad estructural. Aunque se reconoció que la herramienta presentó ideas creativas —como delegar la documentación de instrucciones o emplear una checklist de autoevaluación—, también se identificaron limitaciones importantes. En particular, se observó una falta de soporte numérico sólido en los cálculos propuestos: la hoja de cálculo sugerida no incluía datos reales ni ejemplos completos, y la meta semanal de 2,700 imágenes fue planteada como una proyección ilustrativa, sin justificación cuantitativa. Además, Gemini no consideró las diferencias horarias ni la flexibilidad diaria del equipo, lo que dificultó la implementación de ciertas acciones dependientes de una sincronización estricta, como la revisión de muestras en bloques horarios fijos. A pesar de ello, se implementaron varias de sus recomendaciones: se consolidó la carga de archivos en una sola sesión diaria, se empleó la checklist de revisión individual, se ajustaron metas semanales diferenciadas, se aplicó una revisión focalizada los martes y se envió un reporte diario de errores comunes. También se ejecutó parcialmente la estrategia de metas individualizadas, aunque sin contar con una fórmula completamente desarrollada. En conjunto, Gemini aportó una propuesta extensa, clara y bien estructurada, pero con limitaciones en su adaptación al contexto flexible y asincrónico del equipo.

KPIs obtenidos vs. semana base:

Tabla 10

Comparación de indicadores de desempeño del supervisor: semana base vs. uso de Gemini

Indicador	Valor original	Valor con Gemini	Variación absoluta	Unidad
Tiempo total de supervisión	1390	736	-654	Minutos
Tiempo total de planificación	200	65	-135	Minutos
Eficiencia de planificación	14.39%	8.83%	-5.56%	Porcentaje
Número de muestras revisadas	1440	810	-630	Número de archivos
Tiempo dedicado a revisión diaria	630	362	-268	Minutos
Tiempo promedio por muestra	0.44	0.45	0.01	Minutos/muestra

Nota. Comparación de indicadores clave antes y después de integrar Gemini, mostrando el impacto en tiempos, eficiencia y volumen de trabajo.

Tabla 11

Comparación de indicadores de desempeño del equipo de recolección y etiquetado: semana base vs. uso de Gemini

Tipo	Indicador	Valor original	Valor con Gemini	Variación absoluta	Unidad
Recolección	Total de imágenes recolectadas	2500	2,916	416	Número de imágenes
	Tasa de cumplimiento	100%	100%	0%	Porcentaje
	Tiempo total en recolección	1914	2093	179	Minutos
	Velocidad de recolección	1.35	1.40	0.05	Imágenes/min
	Errores en recolección	35	44	9	Número de errores
	Tasa de error en recolección	1.40%	1.51%	0.11%	Porcentaje
	Precisión de la recolección	98.60%	98.49%	-0.11%	Porcentaje
Etiquetado	Total de imágenes etiquetadas	2500	2,916	416	Número de imágenes
	Tasa de cumplimiento	100%	100%	0%	Porcentaje
	Tiempo total en etiquetado	3273	3686	413	Minutos
	Velocidad de etiquetado	0.77	0.79	0.02	Imágenes/min
	Errores en etiquetado	47	52	5	Número de errores
	Tasa de error en etiquetado	1.88%	1.78%	-0.10%	Porcentaje
	Precisión del etiquetado	98.12%	98.22%	0.10%	Porcentaje
General	Total de miembros	9	9	0	Número de personas
	Productividad general	51.55	51.33	-0.22	Imágenes/hora
	Tiempo total invertido	97.00	113.62	16.62	Horas
	Tiempo total de subida	633	1038	405	Minutos
	Costo por Imagen	0.22	0.19	-0.03	Dólares

Tiempo promedio de subida por archivo	0.13	0.18	0.05	Minutos/imagen
---------------------------------------	------	------	------	----------------

Nota. Comparativa de desempeño antes y después de usar Gemini en recolección y etiquetado, resaltando cambios en volúmenes, velocidades, precisión y costos.

Durante la semana de implementación con Gemini, el tiempo total de supervisión se redujo de manera notable, pasando de 1,390 a 736 minutos, lo que representa un ahorro de 654 minutos. Asimismo, el tiempo dedicado a revisión diaria disminuyó en 268 minutos, aunque el número de muestras revisadas solo se redujo parcialmente (de 1,440 a 810). Sin embargo, la eficiencia de planificación mostró una caída del 14.39% al 8.83%, debido a que la planificación se limitó a 65 minutos en total, sin una sesión estructurada como en semanas anteriores. En cuanto al equipo, se logró un incremento de 416 imágenes tanto en recolección como en etiquetado, alcanzando un total de 2,916 imágenes, lo que refleja una mejora en volumen. La velocidad de recolección subió de 1.35 a 1.40 imágenes por minuto y la de etiquetado de 0.77 a 0.79, con una ligera mejora en la precisión del etiquetado (de 98.12% a 98.22%), aunque se incrementaron los errores en ambas fases. La productividad general se mantuvo estable (51.33 imágenes/hora frente a 51.55 en la semana base), pero el tiempo total invertido aumentó en 16.6 horas y el tiempo de subida de archivos se elevó considerablemente (de 633 a 1,038 minutos), lo que también incrementó el promedio por imagen, a pesar de una reducción positiva en el costo unitario.

Percepción del uso

Al concluir la semana de implementación con Gemini, se aplicaron encuestas estructuradas al supervisor y a los nueve miembros del equipo. Desde la perspectiva del supervisor, las valoraciones fueron mixtas. Si bien se reconoció que las recomendaciones fueron claras y aportaron valor a la gestión, se cuestionó su alineación con los objetivos específicos del proyecto y su capacidad para detectar ineficiencias. La herramienta fue percibida como fácil de usar, con una interfaz comprensible y un proceso de familiarización rápido, aunque la mejora en la calidad del trabajo y la organización general del proyecto no fue plenamente reconocida. Por parte del equipo, las percepciones también fueron diversas: si bien hubo acuerdo general en cuanto a la claridad de las tareas, la utilidad del feedback y la adecuación de la carga de trabajo, solo tres miembros sintieron una mejora clara en su productividad, y varios expresaron neutralidad o desacuerdo respecto a la eficiencia en la gestión del proyecto.

3.4.4 Resultados de la implementación con DeepSeek

Durante la cuarta semana del estudio (del 19 al 23 de mayo de 2025), se implementó la herramienta DeepSeek-V3-0324, que aceptó correctamente todos los archivos enviados (imágenes JPG, texto estructurado y datos de rendimiento). El análisis de los datos tomó apenas 4 segundos, y la generación de recomendaciones se completó en 32 segundos. La respuesta se presentó como un plan de mejora semanal dividido en cuatro bloques claramente definidos, con encabezados por área estratégica, formato conciso y orientado a acciones directas. El texto empleó un estilo técnico-pragmático, con numeración, estimaciones cuantificadas y cálculos específicos en función de los datos reales del equipo, aunque no se incluyó ninguna hoja de cálculo ni ejemplos con fórmulas. A diferencia de otras herramientas, DeepSeek estructuró sus propuestas como un “plan de alto retorno” con impacto proyectado en productividad y eficiencia, sin introducir cambios tecnológicos ni requerir recursos adicionales. Su enfoque estuvo centrado en la redistribución del esfuerzo operativo, la reducción de errores y el ahorro de tiempo para el supervisor, todo dentro del marco real del proyecto.

Resumen de recomendaciones

1. Optimización del proceso:

El sistema detectó que el 85% de los errores se concentraban en solo tres miembros (IDs 104, 105 y 106), mientras que el supervisor invertía 630 minutos semanales en revisar 20 muestras por cada colaborador. Se propuso una revisión focalizada en estos tres miembros (40 muestras diarias) y solo 5 muestras aleatorias del resto, lo cual reduciría el total de muestras revisadas de 1,440 a 600. Además, se sugirió agrupar las imágenes en archivos comprimidos (ZIP) antes de ser subidas, proyectando un ahorro del 50% en el tiempo total de subida.

2. Uso Eficiente de recursos del equipo:

Se identificaron diferencias importantes en la velocidad de trabajo: mientras que los miembros más rápidos (IDs 101–103) alcanzaban hasta 1.5 imágenes/min, los más lentos apenas superaban 0.8 imágenes/min. Asimismo, cuatro miembros (IDs 101, 102, 103 y 109) disponían de una hora diaria no utilizada. DeepSeek recomendó redistribuir la carga: +15 imágenes/día para los miembros rápidos y -10 imágenes/día para los lentos. Esta estrategia permitiría un aumento neto de 225 imágenes semanales sin requerir más tiempo.

3. Mejora de la productividad del equipo:

Se señaló que el principal cuello de botella era el etiquetado, con una velocidad promedio de 0.77 imágenes/min frente a 1.35 en recolección. Además, el 50% de los errores estaban relacionados con categorizaciones incorrectas. Como solución, se recomendó concentrar todo el etiquetado los jueves y viernes, mejorando la consistencia y reduciendo errores en al menos un 10%.

4. Eficiencia del supervisor:

El análisis reveló que 45% del tiempo del supervisor se invertía en revisiones aleatorias, y 180 minutos los viernes se destinaban a correcciones manuales. Se propuso introducir una checklist de errores comunes para retroalimentación estandarizada, lo que ahorraría 20 minutos diarios en comunicación interna (≈ 100 minutos por semana).

5. Plan de acción sugerido por DeepSeek:

DeepSeek propuso un plan práctico basado en datos reales del proyecto, centrado en tres acciones clave con alto retorno esperado:

- Revisión enfocada en los miembros con más errores (IDs 104–106) para reducir el tiempo del supervisor sin comprometer la calidad.
- Redistribución de carga de trabajo, asignando más imágenes a los miembros más eficientes y menos a los que cometían más errores.
- Checklist de errores comunes para estandarizar la retroalimentación y ahorrar tiempo en comunicación.

El modelo estimó como resultado:

- Un aumento del 9% en la producción (hasta 2,725 imágenes),
- Una reducción del 10% en errores, y
- Un ahorro de 6.5 horas semanales del supervisor.

El supervisor valoró la respuesta de DeepSeek con un 4/5 en relevancia, 3/5 en precisión de datos, 4/5 en ejecutabilidad, 3/5 en completitud y 3/5 en claridad estructural. Se reconoció que la herramienta presentó un conjunto de recomendaciones bien enfocadas, claras y alineadas con los objetivos del proyecto, aunque con un nivel de detalle menor al de otras herramientas. Su propuesta semanal fue concisa y directa, sin estructura dividida por días, pero destacó positivamente por priorizar las acciones de mayor impacto esperado y por incluir una sección específica de resultados estimados a una semana vista. Si bien no incorporó archivos de cálculo

ni referencias visuales, su análisis se basó en cifras reales del proyecto y no propuso herramientas adicionales ni cambios estructurales. Además, varias de sus ideas coincidieron con recomendaciones previas de otras herramientas, como la subida por lotes (ChatGPT), la redistribución de carga (Claude), los bloques de etiquetado (ChatGPT) y la checklist de errores comunes (Gemini). Todas las acciones propuestas pudieron implementarse sin dificultades, gracias a que estaban dentro del alcance operativo del equipo y del supervisor.

KPIs obtenidos vs. semana base:

Tabla 12

Comparación de indicadores de desempeño del supervisor: semana base vs. uso de DeepSeek

Indicador	Valor semanal	Valor con DeepSeek	Variación absoluta	Unidad
Tiempo total de supervisión	1390	974	-416	Minutos
Tiempo total de planificación	200	200	0	Minutos
Eficiencia de planificación	14.39%	20.53%	6.15%	Porcentaje
Número de muestras revisadas	1440	600	-840	Número de archivos
Tiempo dedicado a revisión diaria	630	324	-306	Minutos
Tiempo promedio por muestra	0.44	0.54	0.10	Minutos/muestra

Nota. Comparación de indicadores clave antes y después de integrar DeepSeek, mostrando el impacto en tiempos, eficiencia y volumen de trabajo.

Tabla 13

Comparación de indicadores de desempeño del equipo de recolección y etiquetado: semana base vs. uso de DeepSeek

Tipo	Indicador	Valor original	Valor con DeepSeek	Variación absoluta	Unidad
Recolección	Total de imágenes recolectadas	2500	2,780	280	Número de imágenes
	Tasa de cumplimiento	100%	100%	0%	Porcentaje
	Tiempo total en recolección	1914	2121	207	Minutos
	Velocidad de recolección	1.35	1.53	0.18	Imágenes/min
	Errores en recolección	35	36	1	Número de errores
	Tasa de error en recolección	1.40%	1.29%	-0.11%	Porcentaje
	Precisión de la recolección	98.60%	98.71%	0.11%	Porcentaje
Etiquetado	Total de imágenes etiquetadas	2500	2,780	280	Número de imágenes
	Tasa de cumplimiento	100%	100%	0%	Porcentaje
	Tiempo total en etiquetado	3273	3601	328	Minutos
	Velocidad de etiquetado	0.77	0.85	0.08	Imágenes/min
	Errores en etiquetado	47	40	-7	Número de errores

	Tasa de error en etiquetado	1.88%	1.44%	-0.44%	Porcentaje
	Precisión del etiquetado	98.12%	98.56%	0.44%	Porcentaje
General	Total de miembros	9	9	0	Número de personas
	Productividad general	51.55	53.48	1.93	Imágenes/hora
	Tiempo total invertido	97.00	103.97	6.97	Horas
	Tiempo total de subida	633	515.50	-118	Minutos
	Costo por Imagen	0.22	0.20	-0.02	Dólares
	Tiempo promedio de subida por archivo	0.13	0.09	-0.03	Minutos/imagen

Nota. Comparativa de desempeño antes y después de usar DeepSeek en recolección y etiquetado, resaltando cambios en volúmenes, velocidades, precisión y costos.

Durante la semana de implementación con DeepSeek, se observó una mejora significativa en la eficiencia del supervisor. El tiempo total de supervisión se redujo de 1,390 a 974 minutos, principalmente gracias a la revisión enfocada en miembros críticos, lo que también disminuyó el número total de muestras revisadas de 1,440 a 600. Esta estrategia permitió ahorrar 306 minutos en tareas de revisión diaria, al tiempo que aumentó la eficiencia de planificación del 14.39% al 20.53%. Aunque el tiempo promedio por muestra aumentó ligeramente de 0.44 a 0.54 minutos. El nuevo enfoque logró liberar tiempo operativo sin comprometer el control de calidad.

En cuanto al desempeño del equipo, tanto la recolección como el etiquetado de imágenes aumentaron en 280 unidades, alcanzando un total de 2,780 imágenes procesadas en la semana. La tasa de error en etiquetado disminuyó del 1.88% al 1.44%, lo que representa una mejora en la precisión del 0.44%. La velocidad de recolección se redujo de 1.35 a 1.17 imágenes por minuto, mientras que la velocidad de etiquetado mejoró de 0.77 a 0.85 imágenes por minuto, evidenciando una mayor eficiencia en la segunda fase del proceso. la implementación del agrupamiento de archivos y la reducción de errores contribuyeron a una mayor productividad general, que pasó de 51.55 a 53.48 imágenes por hora. Además, el tiempo total de subida de archivos se redujo en 118 minutos, disminuyendo también el costo por imagen de \$0.22 a \$0.20.

Percepción del uso

Desde la perspectiva del supervisor, la evaluación fue mayormente favorable. Se reconoció que el equipo cumplió con las metas dentro de los tiempos previstos, que las recomendaciones fueron coherentes con el contexto real del proyecto y que el uso de DeepSeek tuvo un efecto positivo en la productividad. No obstante, la herramienta no logró identificar ineficiencias previamente no detectadas, y su impacto en la organización general del proyecto fue percibido como limitado. Aun así, se valoró su facilidad de uso, su interfaz clara y el hecho

de no requerir conocimientos técnicos avanzados. En general, la experiencia fue considerada eficiente y se manifestó disposición para reutilizar una herramienta con características similares en el futuro. En cuanto al equipo, las respuestas reflejaron una percepción ampliamente positiva. La mayoría de los miembros coincidió en que las tareas fueron claras desde el inicio, la planificación fue útil y la carga de trabajo adecuada. También destacaron una mejora en su productividad personal y en la organización del flujo de trabajo respecto a semanas anteriores. Sin embargo, algunas respuestas reflejaron experiencias mixtas en relación con la retroalimentación, con dos miembros indicando desacuerdo o neutralidad en este aspecto. A pesar de ello, el equipo percibió el uso de DeepSeek como funcional y beneficioso, especialmente por la claridad de sus recomendaciones y su alineación con los objetivos semanales.

3.4.5 Resultados de la implementación con Infinity

Durante la quinta semana del estudio (del 26 al 30 de mayo de 2025), se utilizó Infinity 1.3.7 Pro para procesar la documentación del proyecto. La herramienta aceptó únicamente archivos en formato Word o PDF, lo que implicó dedicar un estimado de 20 minutos a convertir los datos originales. Una vez cargados, Infinity analizó los insumos en 0 segundos y generó las recomendaciones en 29 segundos. Su respuesta llegó dividida en bloques con encabezados claros para cada área de mejora, integrando cálculos y proyecciones numéricas directamente vinculadas a los valores reales de la semana base, y empleando un lenguaje técnico pero accesible. El tono de la respuesta combinó precisión y cercanía, con formulaciones directas y orientadas a la acción, manteniendo un estilo pragmático que facilitó la comprensión y la implementación inmediata de las sugerencias.

Resumen de recomendaciones

1. Optimización del proceso:

Infinity identificó que el supervisor actualmente dedicaba dos sesiones distintas a la revisión de muestras y la entrega de retroalimentación. Propuso fusionar ambas actividades en un único bloque diario de 120 min, y agrupar la preparación de instrucciones junto con la asignación de tareas en una sola sesión de 110 min los lunes. Con ello, se estima un ahorro de aproximadamente 60 min por semana, liberando tiempo operativo al inicio de cada jornada.

2. Uso eficiente de los recursos del equipo:

Basándose en una productividad actual de 52.1 imágenes/hora por miembro, Infinity elevó el objetivo a 55 imágenes/hora. Manteniendo la misma disponibilidad de cuatro horas diarias durante 5 días, esto se traduce en 220 imágenes adicionales por semana, equivalente a un incremento del 5 % en la productividad total.

3. Mejora de la productividad del equipo:

Para abordar el cuello de botella del etiquetado, la herramienta planteó un sistema de lotes en el que cada colaborador preparase y etiquetase bloques de 100 imágenes antes de realizar la carga. Además, sugirió implementar una sesión semanal de retroalimentación centrada en el análisis de errores comunes y propuestas de mejora. Según Infinity, estas medidas podrían reducir el tiempo de etiquetado en un 10 % y disminuir la tasa de errores en un 15 %.

4. Eficiencia del supervisor:

Infinity recomendó el uso de plantillas predefinidas para las actividades de planificación y revisión, de modo que la planificación total no sobrepase 150 min semanales. Esto representa una reducción del 10 % en la carga de trabajo del supervisor respecto al tiempo invertido en la semana base.

5. Plan de acción sugerido por Infinity:

Para asegurar la implementación efectiva de las recomendaciones de Infinity, se propone un conjunto de acciones prioritarias y de rápida ejecución:

- Implementar el sistema de etiquetado por lotes para mejorar la consistencia y recortar el tiempo de carga en un 10 %.
- Fusionar las sesiones de revisión y retroalimentación en bloques diarios de 120 min para ahorrar 60 min semanales.
- Establecer revisiones semanales centradas en errores comunes y acciones correctivas, con el objetivo de bajar la tasa de fallos al 1.60 %.

El supervisor valoró la respuesta de Infinity con un 3/5 en relevancia, 2/5 en precisión de datos, 3/5 en ejecutabilidad, 2/5 en completitud y 3/5 en claridad estructural. Aunque las recomendaciones fueron prácticas y directas, la respuesta fue la más breve entre todas las herramientas analizadas y presentó debilidades importantes en el análisis técnico. La mayoría de las acciones carecían de cálculos detallados o fundamentos cuantitativos precisos, y se basaban en estimaciones generales sin justificaciones específicas ni vinculación clara con los datos reales

del proyecto. El plan de acción no estuvo organizado por días, como en otros casos, sino por orden de prioridad, lo cual facilitó su lectura, pero limitó su aplicabilidad dentro de un cronograma operativo. A pesar de ello, se implementaron varias de sus sugerencias: se adoptó la estrategia de etiquetado por lotes, se consolidaron las sesiones de retroalimentación en bloques únicos diarios, y se introdujo una sesión semanal para discutir errores comunes. También se ajustaron las metas individuales a 55 imágenes por hora, logrando un incremento en la producción semanal. Sin embargo, algunas propuestas mostraron inconsistencias, como el uso de bloques de 100 imágenes sin una lógica clara con respecto al volumen real de trabajo diario por miembro, y la falta de detalles operativos en acciones como la fusión de tareas de planificación. En conjunto, Infinity ofreció ideas útiles y aplicables en poco tiempo, pero con una propuesta menos robusta en términos de análisis técnico y profundidad comparativa.

KPIs obtenidos vs. semana base:

Tabla 14

Comparación de indicadores de desempeño del supervisor: semana base vs. uso de Infinity

Indicador	Valor semanal	Valor con Infinity	Variación absoluta	Unidad
Tiempo total de supervisión	1390	952	-438	Minutos
Tiempo total de planificación	200	182	-18	Minutos
Eficiencia de planificación	14.39	19.12	4.73	Porcentaje
Número de muestras revisadas	1440	1078	-362	Número de archivos
Tiempo dedicado a revisión diaria	630	480	-150	Minutos
Tiempo promedio por muestra	0.44	0.45	0.01	Minutos/muestra

Nota. Comparación de indicadores clave antes y después de integrar Infinity, mostrando el impacto en tiempos, eficiencia y volumen de trabajo.

Tabla 15

Comparación de indicadores de desempeño del equipo de recolección y etiquetado: semana base vs. uso de Infinity

Tipo	Indicador	Valor original	Valor con Infinity	Variación absoluta	Unidad
Recolección	Total de imágenes recolectadas	2500	2,736	236	Número de imágenes
	Tasa de cumplimiento	100%	100%	0%	Porcentaje
	Tiempo total en recolección	1914	2093	179	Minutos

	Velocidad de recolección	1.35	1.36	0.01	Imágenes/min
	Errores en recolección	35	41	6	Número de errores
	Tasa de error en recolección	1.40%	1.50%	0.10%	Porcentaje
	Precisión de la recolección	98.60%	98.50%	-0.10%	Porcentaje
Etiquetado	Total de imágenes etiquetadas	2500	2,736	236	Número de imágenes
	Tasa de cumplimiento	100%	100%	0%	Porcentaje
	Tiempo total en etiquetado	3273	3421	148	Minutos
	Velocidad de etiquetado	0.77	0.80	0.03	Imágenes/min
	Errores en etiquetado	47	56	9	Número de errores
	Tasa de error en etiquetado	1.88%	2.05%	0.17%	Porcentaje
	Precisión del etiquetado	98.12%	97.95%	-0.17%	Porcentaje
General	Total de miembros	9	9	0	Número de personas
	Productividad general	51.55	52.48	0.94	Imágenes/hora
	Tiempo total invertido	97.00	104.26	7.26	Horas
	Tiempo total de subida	633	742.00	109	Minutos
	Costo por Imagen	0.22	0.20	-0.02	Dólares
	Tiempo promedio de subida por archivo	0.13	0.14	0.01	Minutos/imagen

Nota. Comparativa de desempeño antes y después de usar Infinity en recolección y etiquetado, resaltando cambios en volúmenes, velocidades, precisión y costos.

Durante la semana de implementación con Infinity, el tiempo total de supervisión disminuyó de 1,390 a 952 minutos, representando un ahorro de 438 minutos gracias a la consolidación de bloques diarios de revisión y retroalimentación. La planificación semanal también se optimizó, reduciéndose en 18 minutos, lo que elevó la eficiencia de planificación del 14.39 % al 19.12 %. Aunque el número de muestras revisadas bajó de 1,440 a 1,078, el tiempo promedio por muestra se mantuvo prácticamente constante, pasando de 0.44 a 0.45 minutos.

En cuanto al desempeño del equipo, se registró un incremento de 236 imágenes en recolección y etiquetado, alcanzando un total de 2,736 imágenes en ambas fases. La productividad general creció ligeramente, de 51.55 a 52.48 imágenes por hora. La velocidad de etiquetado también mejoró levemente (de 0.77 a 0.80 imágenes/min), mientras que la velocidad de recolección se mantuvo estable. Sin embargo, el número total de errores en ambas actividades aumentó, y se evidenció un leve incremento en la tasa de error, particularmente en el etiquetado (de 1.88 % a 2.05 %), lo que provocó una disminución en la precisión general. El tiempo total invertido por el equipo aumentó en 7.26 horas, acompañado de un alza en el tiempo total de subida (de 633 a 742 minutos), lo cual afectó ligeramente el tiempo promedio por archivo (de 0.13 a 0.14 minutos). A pesar de ello, el costo por imagen se redujo a 0.20 USD, manteniéndose dentro del rango esperado.

Percepción del uso

Desde la perspectiva del supervisor, las respuestas fueron mayoritariamente neutrales. Aunque se reconoció que el equipo cumplió sus metas dentro del tiempo previsto y que las recomendaciones fueron claras y adecuadas al contexto real, se observaron limitaciones importantes. La herramienta no logró evidenciar mejoras significativas en la organización o el seguimiento del proyecto, ni permitió detectar ineficiencias que no hubieran sido identificadas previamente. Asimismo, se percibió poco impacto directo en la productividad y cierta desconexión entre las recomendaciones y los objetivos específicos del proyecto. A pesar de estas limitaciones, se destacó positivamente la facilidad de uso, la interfaz intuitiva y el hecho de que no se requirieron conocimientos técnicos avanzados para su implementación. Por parte del equipo, las percepciones fueron más equilibradas. La mayoría de los miembros coincidió en que las tareas estuvieron claras desde el inicio, que la planificación diaria fue útil y que la carga de trabajo fue adecuada y manejable. También se observó una percepción general favorable en cuanto a la organización del flujo de trabajo y la productividad personal. No obstante, se evidenció cierta variabilidad en la retroalimentación recibida, con uno de los miembros en desacuerdo y otros mostrando respuestas más neutras. En conjunto, Infinity fue valorada como una herramienta funcional y práctica, aunque con un impacto más limitado en comparación con otras implementadas anteriormente.

3.5 SÍNTESIS COMPARATIVA ENTRE HERRAMIENTAS

3.5.1 Tabla consolidada de desempeño por herramienta y KPI

La siguiente tabla resume los cambios en la carga de trabajo y eficiencia del supervisor al implementar cada herramienta GenAI, comparando métricas clave como tiempo total de supervisión, planificación, eficiencia organizativa y esfuerzo de revisión. Permite observar el impacto directo de cada herramienta sobre la gestión operativa del rol supervisor.

Tabla 16

Comparación de indicadores de desempeño del supervisor: semana base vs. herramientas de GenAI

Indicador	Valor original	Valor con ChatGPT	Valor con Claude	Valor con Gemini	Valor con DeepSeek	Valor con Infinity	Unidad
Tiempo total de supervisión	1390	980	1006	736	974	952	Minutos
Tiempo total de planificación	200	200	200	65	200	182	Minutos
Eficiencia de planificación	14.39	20.41	19.88	8.83	20.53	19.12	Porcentaje

Número de muestras revisadas	1440	720	540	810	600	1078	Número de archivos
Tiempo dedicado a revisión diaria	630	490	561	362	324	480	Minutos
Tiempo promedio por muestra	0.44	0.68	1.04	0.45	0.54	0.45	Minutos/muestra

Nota. Comparativa de desempeño del supervisor usando distintas herramientas GenAI, mostrando variaciones en tiempos, eficiencia de planificación y número de muestras revisadas.

La tabla de abajo presenta los principales indicadores de desempeño del equipo durante las 5 semanas de implementación, en comparación con la semana base. Incluye métricas de volumen de imágenes, velocidad, errores, precisión, cumplimiento de metas, tiempos invertidos, y eficiencia económica, ofreciendo una visión integral del rendimiento del equipo bajo cada intervención tecnológica.

Tabla 17

Comparación de indicadores de desempeño del equipo de recolección y etiquetado: semana base vs. herramientas de GenAI

Tipo	Indicador	Valor original	Valor con ChatGPT	Valor con Claude	Valor con Gemini	Valor con DeepSeek	Valor con Infinity	Unidad
Recolección	Total de imágenes recolectadas	2500	2,700	2,632	2,916	2,780	2,736	Número de imágenes
	Tasa de cumplimiento	100%	100%	100%	100%	100%	100%	Porcentaje
	Tiempo total en recolección	1914	2275	1756	2093	2121	2093	Minutos
	Velocidad de recolección	1.35	1.20	1.58	1.40	1.53	1.36	Imágenes/min
	Errores en recolección	35	41	36	44	36	41	Número de errores
	Tasa de error en recolección	1.40%	1.52%	1.37%	1.51%	1.29%	1.50%	Porcentaje
	Precisión de la recolección	98.60%	98.48%	98.63%	98.49%	98.71%	98.50%	Porcentaje
Etiquetado	Total de imágenes etiquetadas	2500	2,700	2,632	2,916	2,780	2,736	Número de imágenes
	Tasa de cumplimiento	100%	100%	100%	100%	100%	100%	Porcentaje
	Tiempo total en etiquetado	3273	3138	3440	3686	3601	3421	Minutos
	Velocidad de etiquetado	0.77	0.86	0.77	0.79	0.85	0.80	Imágenes/min
	Errores en etiquetado	47	25	54	52	40	56	Número de errores
	Tasa de error en etiquetado	1.88%	0.93%	2.05%	1.78%	1.44%	2.05%	Porcentaje
	Precisión del etiquetado	98.12%	99.07%	97.95%	98.22%	98.56%	97.95%	Porcentaje
General	Total de miembros	9	9	9	9	9	9	Número de personas
	Productividad general	51.55	56.47	50.98	51.33	53.48	52.48	Imágenes/hora
	Tiempo total invertido	97.00	95.63	103.25	113.62	103.97	104.26	Horas
	Tiempo total de subida	633	325	999	1038	515	742	Minutos
	Costo por Imagen	0.22	0.20	0.21	0.19	0.20	0.20	Dólares
	Tiempo promedio de subida por archivo	0.13	0.06	0.19	0.18	0.09	0.14	Minutos/imagen

Nota. Comparativa del desempeño del equipo con varias herramientas GenAI, destacando variaciones en volúmenes, tiempos, errores y productividad.

CAPÍTULO IV – ANÁLISIS COMPARATIVO Y TOMA DE DECISIONES

4.1 RESUMEN DE LOS PRINCIPALES HALLAZGOS

4.1.1 Comparación cuantitativa de KPIs

En esta sección se presenta un resumen de cómo cada herramienta de GenAI influyó en los indicadores clave de desempeño (KPIs) del proyecto, tanto desde la perspectiva del supervisor como del equipo de recolección y etiquetado. Partiendo de las Tablas Indicadores del supervisor y equipo, se calculó para cada KPI la variación absoluta y porcentual respecto a la semana base, de modo que sea posible identificar de un vistazo qué solución generó mayores mejoras o retrocesos. Para facilitar un análisis más eficiente, se ha seleccionado los KPIs más relevantes —tanto del rol de supervisión como del desempeño del equipo— aquellos que reflejan cambios significativos o que son clave para la gestión; el resto de los indicadores, por su menor impacto o por mostrar variaciones mínimas o nulas durante las pruebas, no se han considerado en este resumen.

Análisis cuantitativo de los KPIs clave del supervisor

A continuación, se presenta una tabla con los valores absolutos de cada KPI de supervisor frente a la semana base, y se muestran, para cada indicador importante, gráficos de barras que ilustran la variación porcentual con respecto al valor base.

Tabla 18

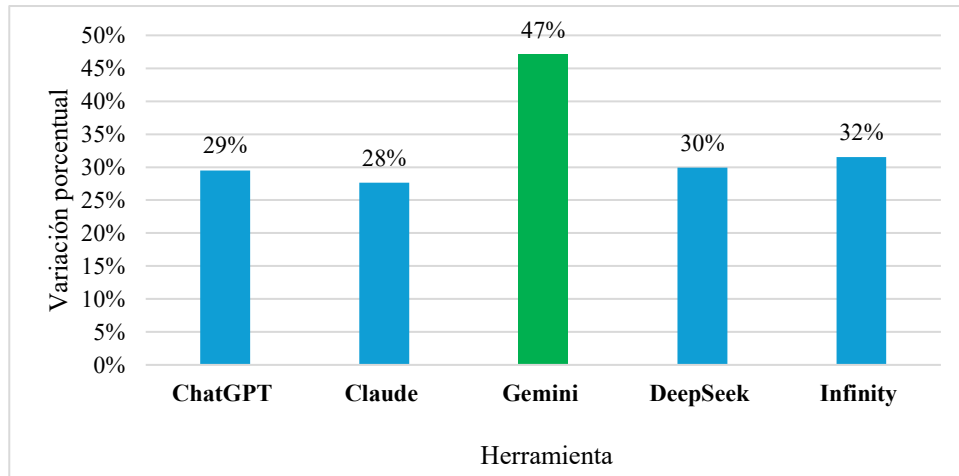
Valores absolutos de los KPIs clave de supervisión: semana base vs. herramientas de GenAI

Indicador	Valor original	Valor absoluto con ChatGPT	Valor absoluto con Claude	Valor absoluto con Gemini	Valor absoluto con DeepSeek	Valor absoluto con Infinity	Unidad
Tiempo total de supervisión	1390	-410	-384	-654	-416	-438	Minutos
Eficiencia de planificación	14.39%	6.02%	5.49%	-5.56%	6.15%	4.73%	Porcentaje
Número de muestras revisadas	1440	-720	-900	-630	-840	-362	Número de archivos
Tiempo dedicado a revisión diaria	630	-140	-69	-268	-306	-150	Minutos

Nota. Los valores mostrados representan la diferencia absoluta de cada KPI tras la implementación de la herramienta GenAI respecto a la semana base (nuevo valor – valor base).

Figura 3

Variación porcentual del tiempo total de supervisión

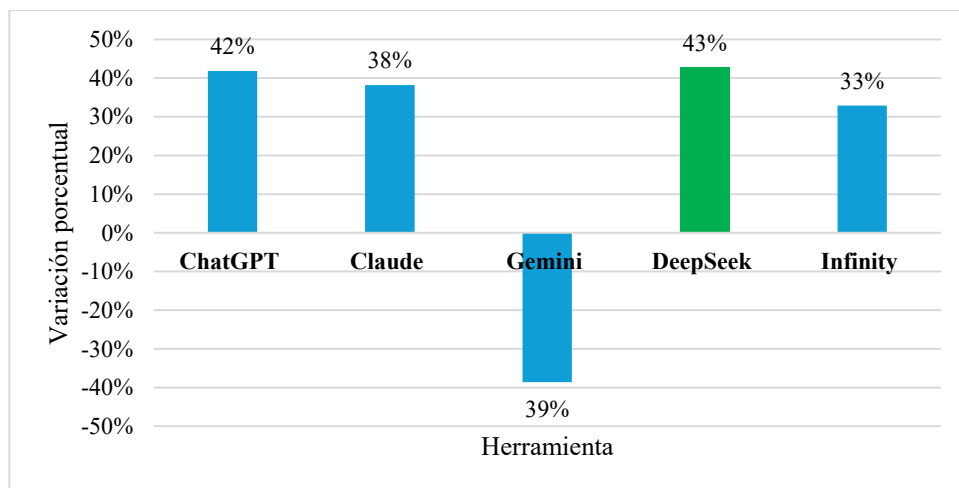


Nota. $\Delta\% = (\text{valor herramienta} - \text{valor base}) / \text{valor base} \times 100$. Los valores indican menor tiempo de supervisión.

Gemini demostró ser la opción más eficaz, reduciendo el tiempo de supervisión en 654 minutos, lo que representa una disminución del 47 % respecto a la semana base. Esta mejora significativa libera casi 11 horas de dedicación del supervisor, permitiendo redirigir ese tiempo a tareas de mayor valor estratégico y fortaleciendo la agilidad en la gestión del proyecto.

Figura 4

Variación porcentual de la eficiencia de planificación

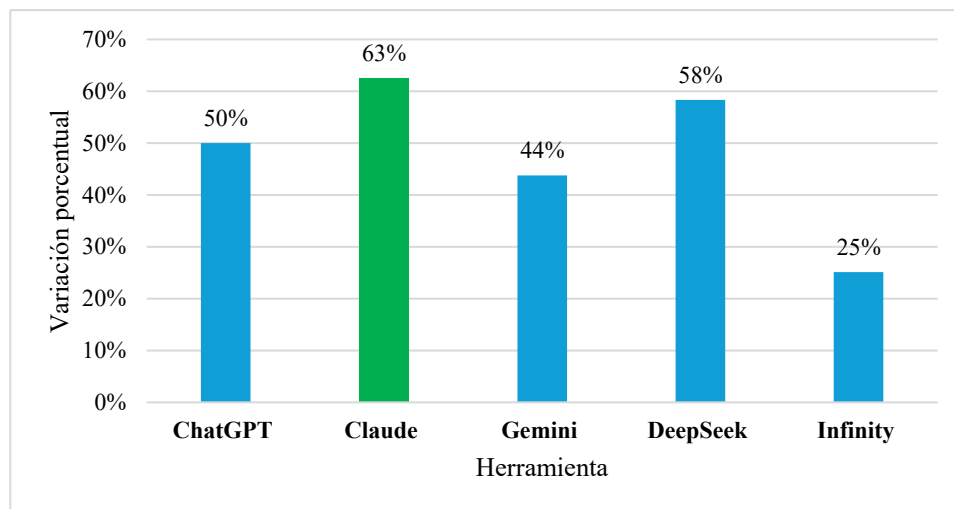


Nota. $\Delta\% = (\text{valor herramienta} - \text{valor base}) / \text{valor base} \times 100$. Los valores positivos indican mejoras.

DeepSeek se alza como la herramienta más eficaz para mejorar la eficiencia de planificación, incrementando este KPI en 6,15 puntos porcentuales, al pasar de un 14,39 % en la semana base a un 20,53 % durante su uso. Esto supone un aumento relativo del 43 %, el mayor registrado entre todas las soluciones analizadas. Gracias a esta mejora, la dedicación del supervisor a tareas de planificación rinde casi la mitad más.

Figura 5

Variación porcentual del número de muestras revisadas

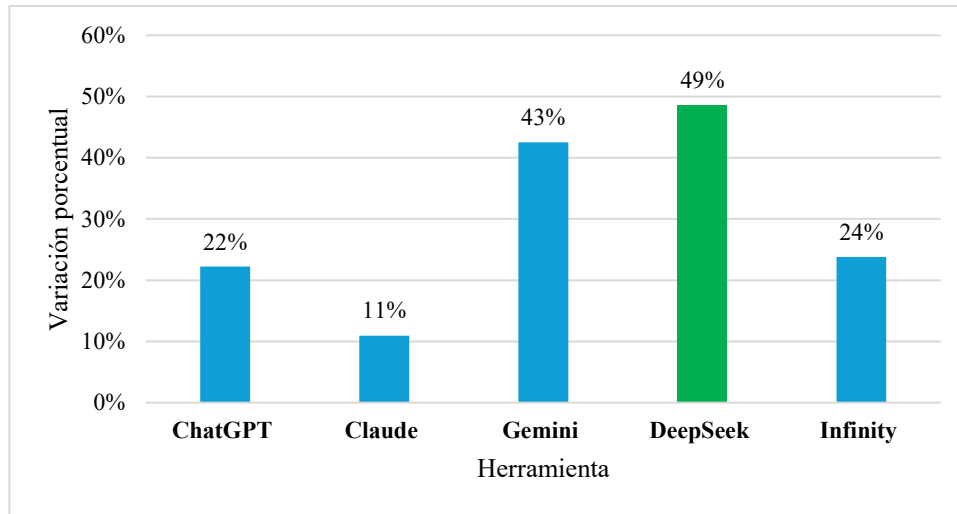


Nota. $\Delta \% = (\text{valor herramienta} - \text{valor base}) / \text{valor base} \times 100$. Los valores indican una disminución en el número de muestras revisadas.

En el KPI Número de muestras revisadas, Claude se impone como la herramienta más efectiva para aliviar la carga de revisión, con una disminución del 63 % (de 1 440 a 540 muestras). Este descenso representa la mayor reducción en el volumen de trabajo de revisión, liberando al supervisor de revisar casi dos tercios de las muestras habituales.

Figura 6

Variación porcentual del tiempo dedicado a revisión diaria



Nota. $\Delta \% = (\text{valor herramienta} - \text{valor base}) / \text{valor base} \times 100$. Los valores indican reducción en el tiempo invertido en revisar muestras.

En el KPI Tiempo dedicado a revisión diaria, DeepSeek fue la herramienta más efectiva, reduciendo este tiempo en 306 minutos, lo que representa una disminución relativa del 49 % respecto a la semana base. Esta importante reducción libera más de 5 horas diarias del supervisor.

Análisis cuantitativo de los KPIs clave del equipo

A continuación, se muestra la tabla de valores absolutos de los KPIs clave del equipo, seguida de gráficos de barras que ilustran la variación porcentual de cada indicador frente al valor base. Cada figura presentará el $\Delta \%$ con un código de color (verde) para facilitar la identificación rápida de la herramienta que mejor optimiza cada métrica de equipo.

Tabla 19

Valores absolutos de los KPIs clave del equipo: semana base vs. herramientas de GenAI

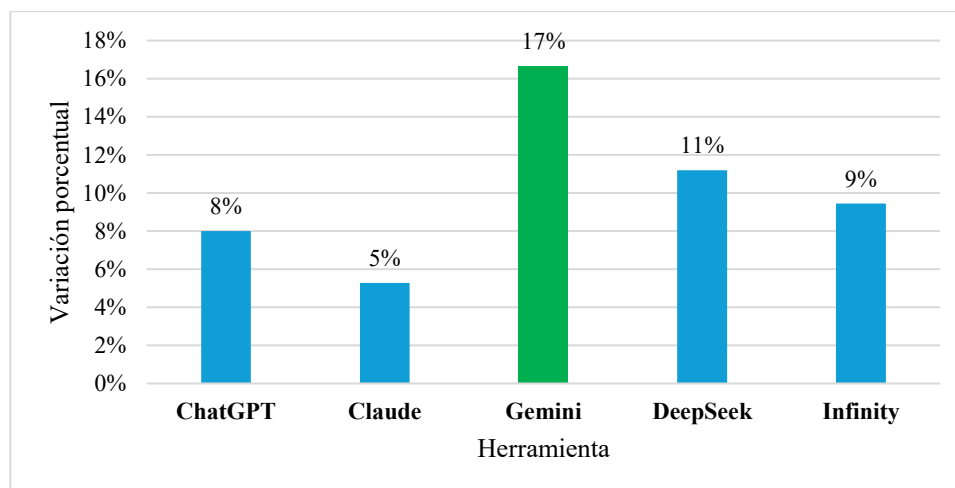
Tipo	Indicador	Valor original	Variación absoluta ChatGPT	Variación absoluta Claude	Variación absoluta Gemini	Variación absoluta DeepSeek	Variación absoluta Infinity	Unidad
Recolección	Total de imágenes recolectadas	2500	200	132	416	280	236	Número de imágenes
	Velocidad de recolección	1.35	-0.15	0.23	0.05	0.18	0.01	Imágenes/min
	Tasa de error en recolección	1.40%	0.12%	-0.03%	0.11%	-0.11%	0.10%	Porcentaje

Etiquetado	Total de imágenes etiquetadas	2500	200	132	416	280	236	Número de imágenes
	Velocidad de etiquetado	0.77	0.09	0.00	0.02	0.08	0.03	Imágenes/min
	Tasa de error en etiquetado	1.88%	-0.95%	0.17%	-0.10%	-0.44%	0.17%	Porcentaje
General	Productividad general	51.55	4.92	-0.56	-0.22	1.93	0.94	Imágenes/hora
	Tiempo total de subida	633	-308	366	405	-118	109	Minutos
	Costo por Imagen	0.22	-0.02	-0.01	-0.03	-0.02	-0.02	Dólares

Nota. Los valores mostrados representan la diferencia absoluta de cada KPI tras la implementación de la herramienta GenAI respecto a la semana base (nuevo valor – valor base).

Figura 7

Variación porcentual del total de imágenes recolectadas

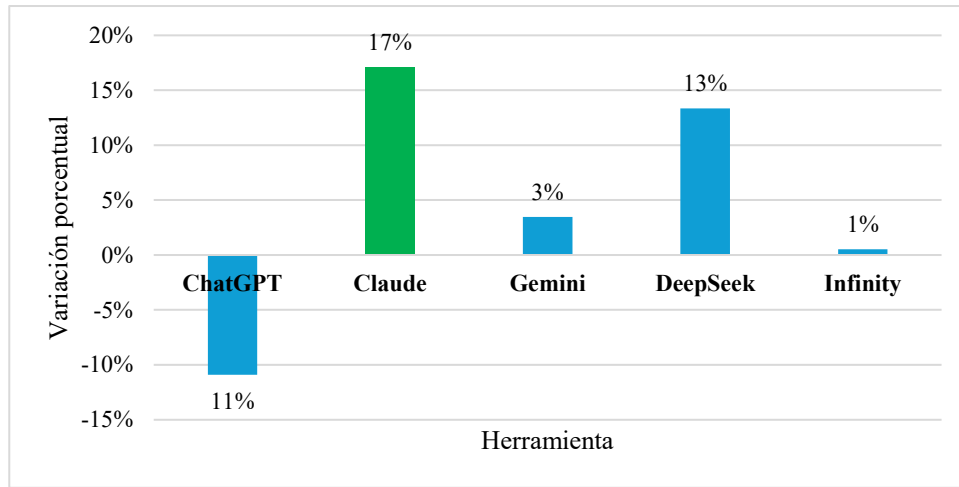


Nota. $\Delta \% = (\text{valor herramienta} - \text{valor base}) / \text{valor base} \times 100$. Los valores positivos indican incrementos en el volumen de imágenes recolectadas.

En el KPI Total de imágenes recolectadas, Gemini obtuvo el mejor desempeño, aumentando el volumen en 416 imágenes, lo que representa un crecimiento del 17 % respecto a la semana base. Esta mejora se traduce en una capacidad significativamente mayor para alcanzar los objetivos de recolección sin requerir horas adicionales.

Figura 8

Variación porcentual de la velocidad de recolección

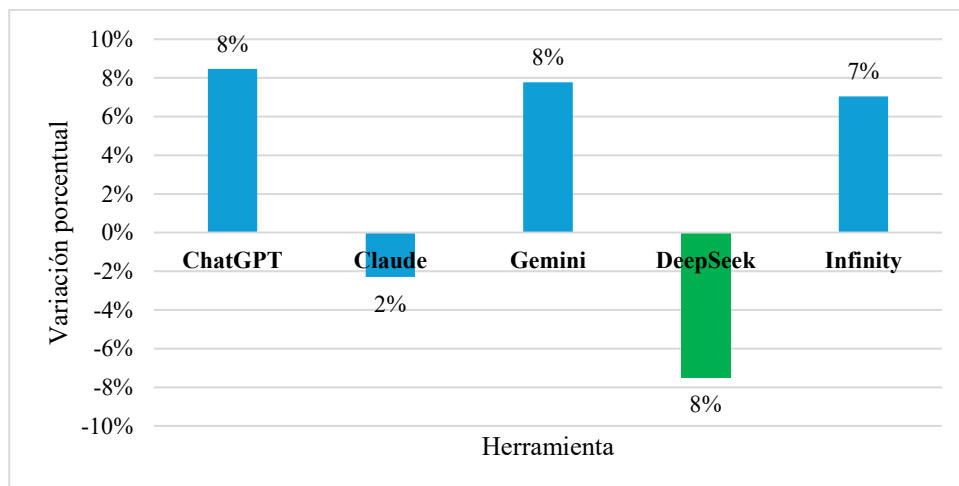


Nota. $\Delta \% = (\text{valor herramienta} - \text{valor base}) / \text{valor base} \times 100$. Los valores positivos indican incrementos en la velocidad de recolección, y los negativos disminuciones.

En el KPI Velocidad de recolección, Claude destacó con un incremento del 17 %, equivalente a sumar 0,23 imágenes por minuto al ritmo de trabajo, lo que mejora significativamente la capacidad de procesamiento sin aumentar el tiempo disponible.

Figura 9

Variación porcentual de la tasa de error en recolección

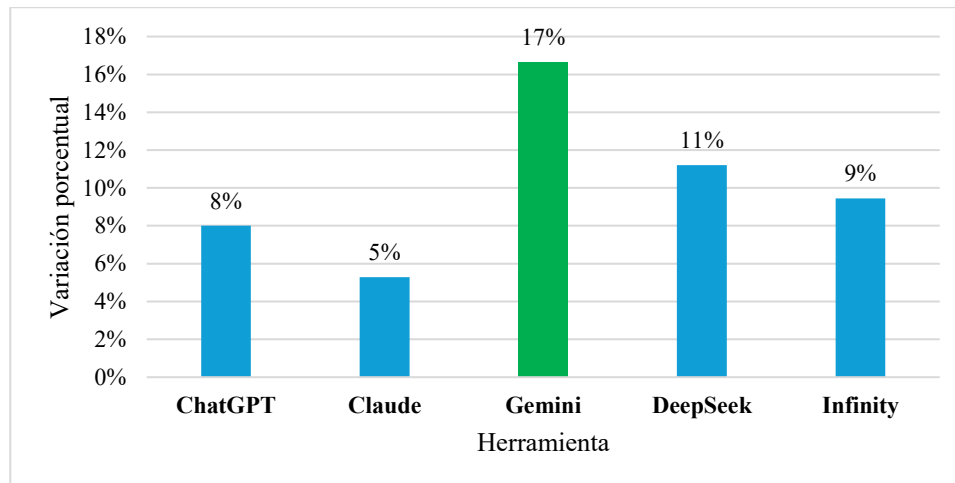


Nota. $\Delta \% = (\text{valor herramienta} - \text{valor base}) / \text{valor base} \times 100$. Valores negativos indican reducción en la tasa de error (mejora), mientras que los positivos señalan aumentos en errores.

En el KPI Tasa de error en recolección, DeepSeek fue la herramienta más efectiva al reducir la tasa de error en 0,11 puntos porcentuales, lo que equivale a una disminución relativa del 8 %. Esta mejora contribuye a una mayor calidad de los datos capturados, minimizando retrabajos y fortaleciendo la fiabilidad del proceso de recolección.

Figura 10

Variación porcentual del total de imágenes etiquetadas

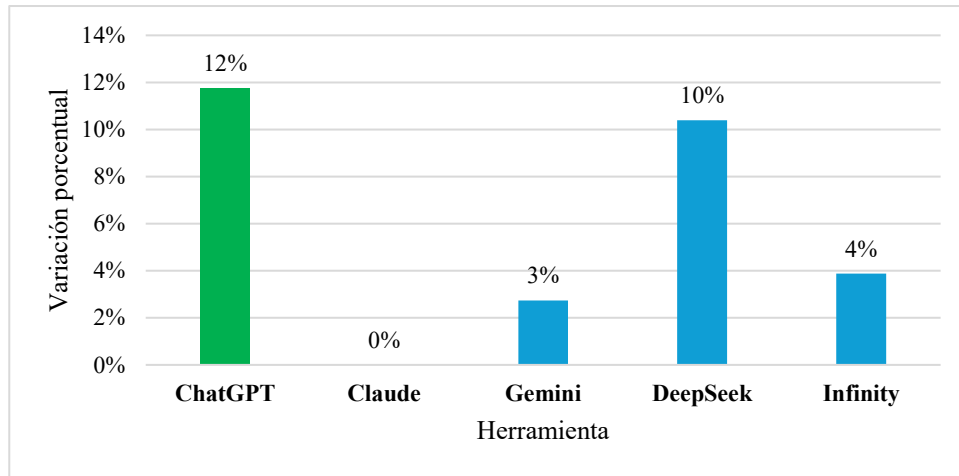


Nota. $\Delta \% = (\text{valor herramienta} - \text{valor base}) / \text{valor base} \times 100$. Los valores positivos indican incrementos en el volumen de imágenes etiquetadas.

Para el KPI Total de imágenes etiquetadas, Gemini vuelve a destacarse como la opción más potente, aumentando el volumen en 416 imágenes—un alza del 17 % respecto a la semana base. Este incremento refuerza la capacidad del equipo para procesar mayores cantidades sin requerir más horas, mejorando la eficiencia operativa global.

Figura 11

Variación porcentual de la velocidad de etiquetado

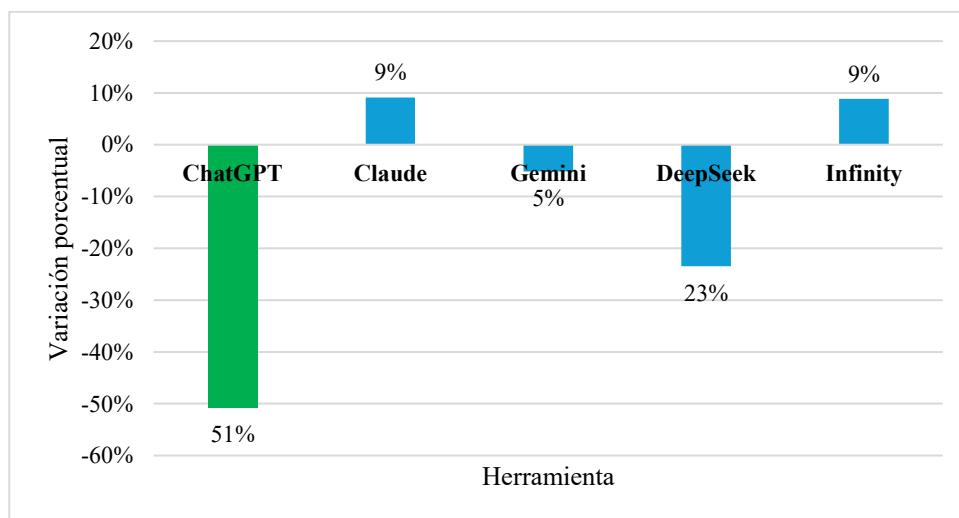


Nota. $\Delta \% = (\text{valor herramienta} - \text{valor base}) / \text{valor base} \times 100$. Los valores positivos indican incrementos en la velocidad de etiquetado.

En el KPI Velocidad de etiquetado, ChatGPT resultó la herramienta más eficaz, aumentando el ritmo un 12 %—esto equivale a procesar aproximadamente 0,09 imágenes/minutos adicionales—lo que mejora la capacidad del equipo para completar anotaciones con mayor rapidez y cumplir objetivos en menor tiempo.

Figura 12

Variación porcentual de la tasa de error en etiquetado

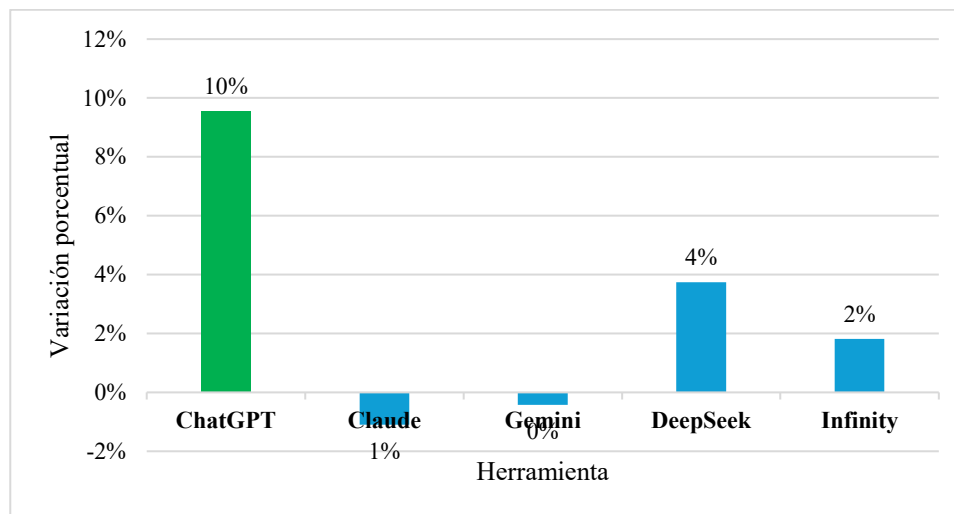


Nota. $\Delta \% = (\text{valor herramienta} - \text{valor base}) / \text{valor base} \times 100$. Valores negativos indican reducción en la tasa de error (mejora), y los positivos aumentos en errores.

En el KPI Tasa de error en etiquetado, ChatGPT demostró ser la herramienta más eficiente, reduciendo los errores en 0,95 puntos porcentuales (de 1,88 % a 0,93 %), lo que equivale a una disminución relativa del 51 %. Esta fuerte reducción mejora la calidad de las anotaciones y minimiza la necesidad de re-trabajo, elevando la fiabilidad del proceso de etiquetado.

Figura 13

Variación porcentual de la productividad general

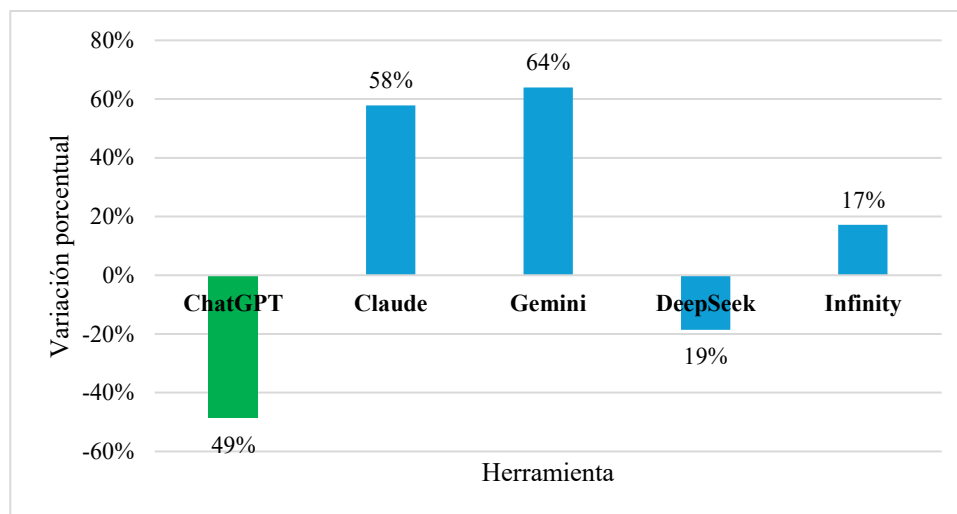


Nota. $\Delta \% = (\text{valor herramienta} - \text{valor base}) / \text{valor base} \times 100$. Los valores positivos indican incrementos en la productividad (imágenes por hora).

En el KPI Productividad general, ChatGPT destacó como la herramienta más eficaz, elevando la productividad en 4,92 imágenes/hora, lo que supone un aumento del 10 % respecto a la semana base. Este avance mejora el rendimiento del equipo sin necesidad de horas adicionales, optimizando el uso del tiempo disponible.

Figura 14

Variación porcentual del tiempo total de subida

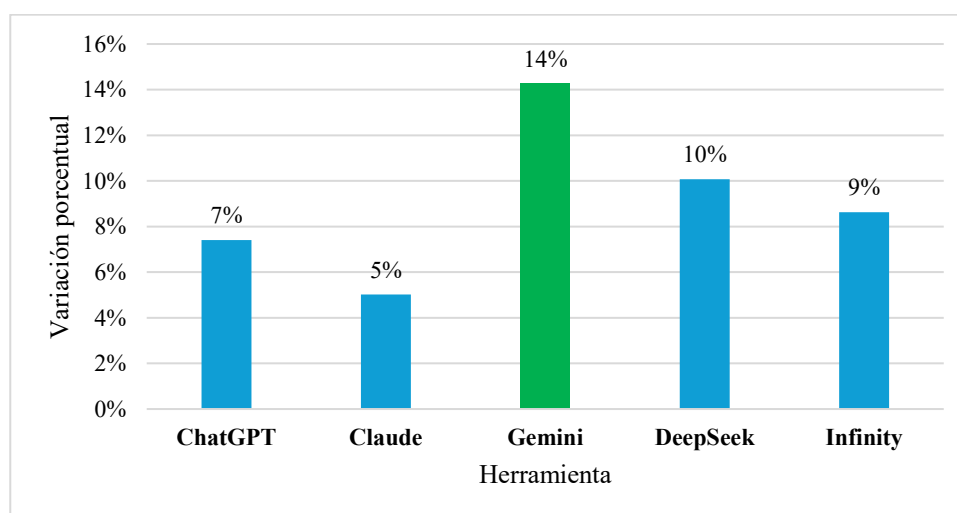


Nota. $\Delta \% = (\text{valor herramienta} - \text{valor base}) / \text{valor base} \times 100$. Valores negativos indican reducción en el tiempo de subida (mejora), positivos incrementos.

En el KPI Tiempo total de subida, ChatGPT se destaca como la herramienta más eficiente, reduciendo el tiempo de carga en 308 minutos, lo que equivale a una disminución del 49 % respecto a la semana base. Esta mejora libera más de 5 horas de tareas de subida, acelerando significativamente el flujo de trabajo y reduciendo cuellos de botella en la gestión de archivos.

Figura 15

Variación porcentual del costo por imagen



Nota. $\Delta\% = (\text{valor herramienta} - \text{valor base}) / \text{valor base} \times 100$. Los valores indican reducción en el costo por imagen (mejora).

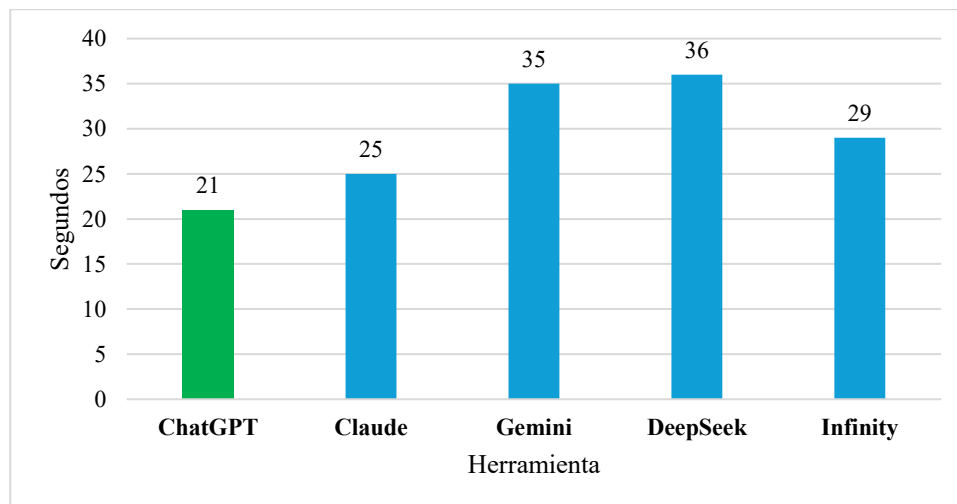
En el KPI Costo por imagen, Gemini resultó la opción más económica, reduciendo el costo en \$0,03 por imagen, lo que equivale a una disminución relativa del 14 % respecto a la semana base. Esta reducción optimiza el presupuesto del proyecto sin sacrificar calidad ni eficiencia operativa.

Latencia de respuesta de las herramientas GenAI

A continuación, se incorpora el KPI cuantitativo que mide la latencia de respuesta de cada herramienta GenAI, es decir, el tiempo total (en segundos) que tarda en procesar los datos del proyecto y generar sus recomendaciones. Este indicador resulta clave para evaluar la rapidez de adopción de cada solución en un entorno operativo real.

Figura 16

Tiempo de respuesta de las herramientas GenAI



Nota. Valores en segundos que representan el tiempo transcurrido desde la carga de datos hasta la generación de recomendaciones por cada herramienta.

En términos de latencia de respuesta, ChatGPT fue la más ágil, completando el análisis y entrega de recomendaciones en solamente 21 segundos, seguida por Claude (25 s) e Infinity (29 s). Por el contrario, Gemini y DeepSeek mostraron tiempos mayores (35 s y 36 s respectivamente).

4.1.2 Análisis cualitativo de las recomendaciones

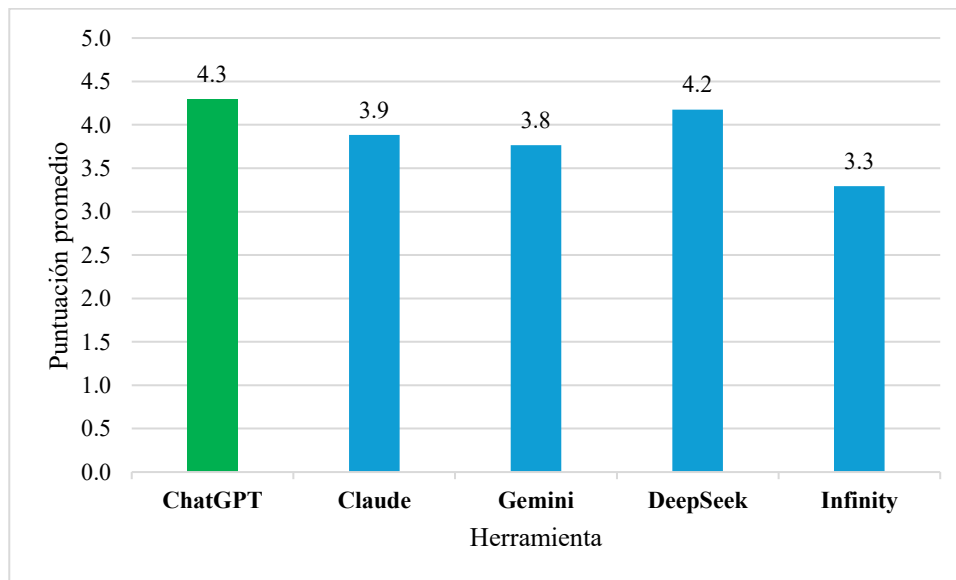
En esta sección se profundiza la percepción y valoración de las sugerencias generadas por cada herramienta GenAI, complementando los resultados cuantitativos con la experiencia de los usuarios. Para ello, se presenta la puntuación media global que el supervisor otorgó en su encuesta de satisfacción, así como la puntuación media global recabada entre los miembros del equipo de recolección y etiquetado. A continuación, se detalla la evaluación técnica de cada herramienta según cinco criterios específicos (relevancia, precisión de datos, ejecutabilidad, completitud y claridad/estructura), reflejando cómo se compara la calidad de sus respuestas más allá de los KPIs operativos.

Satisfacción general del supervisor

Aquí se muestra la puntuación media que el supervisor asignó a cada GenAI al término de la semana de implementación, en una escala de 1 a 5, evaluando relevancia, claridad y aplicabilidad de sus recomendaciones.

Figura 17

Puntuación media global del supervisor



Nota. Cada valor muestra el promedio (escala 1–5) otorgado por el supervisor en su encuesta de satisfacción sobre relevancia, claridad y aplicabilidad de las recomendaciones de cada herramienta GenAI al final de la semana de implementación.

ChatGPT lidera con una puntuación media de 4,3, reflejando la mejor valoración global del supervisor en términos de utilidad y calidad de sus sugerencias. Le sigue muy de cerca

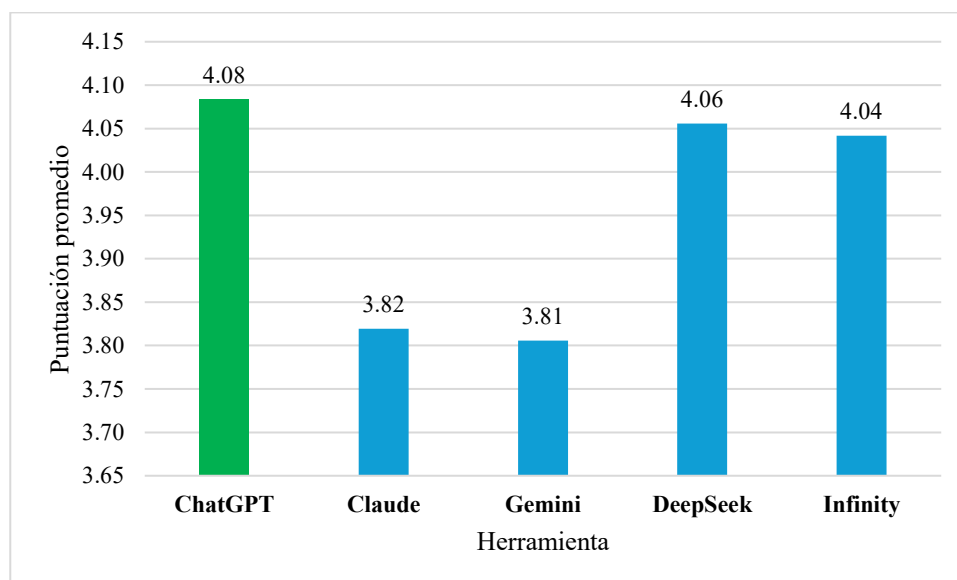
DeepSeek con 4,2, mientras que Infinity obtiene el resultado más bajo (3,3), lo que sugiere que sus recomendaciones podrían requerir ajustes en profundidad o formato para mejorar su adoptabilidad.

Satisfacción media del equipo

A continuación, se presenta la valoración promedio del equipo de recolección y etiquetado, basada en criterios de utilidad práctica, claridad de las instrucciones y facilidad de implementación de las sugerencias recibidas.

Figura 18

Puntuación media global del equipo



Nota. Cada valor refleja el promedio (escala 1–5) de las respuestas del equipo sobre la utilidad práctica, la claridad y la aplicabilidad de cada GenAI al final de la semana.

ChatGPT encabeza la satisfacción del equipo con una puntuación media de 4,08, demostrando que sus recomendaciones fueron las más valoradas por los colaboradores en términos de claridad y facilidad de implementación. Le siguen DeepSeek (4,06) e Infinity (4,04), mientras que Gemini (3,81) y Claude (3,82) reciben calificaciones ligeramente inferiores.

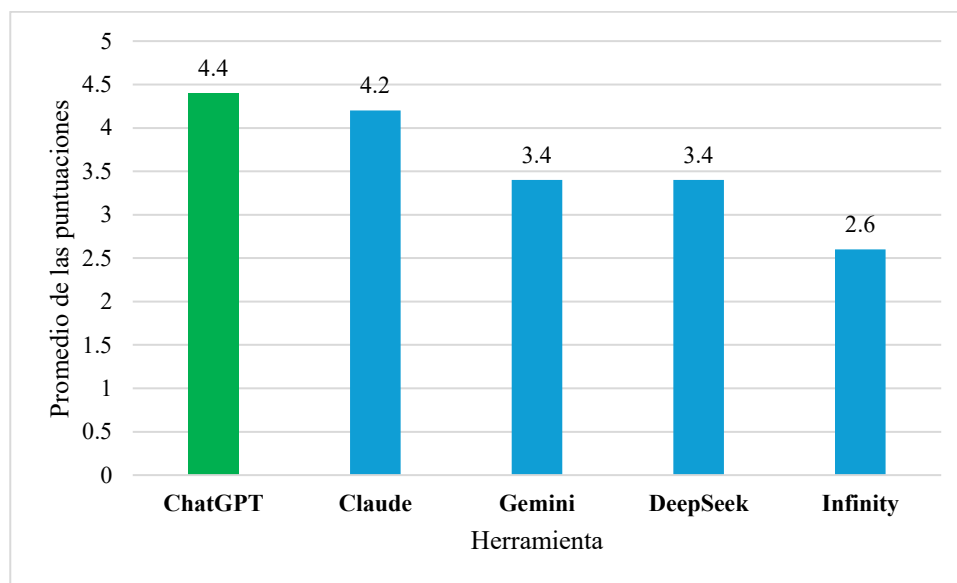
Evaluación de criterios técnicos por herramienta

Por último, se desglosa la calificación técnica de cada GenAI según relevancia, precisión de datos, ejecutabilidad, completitud y claridad/estructura, basándose en la opinión del

supervisor —como gerente de equipo y responsable de implementar las recomendaciones— para comparar la calidad de sus respuestas más allá de los KPIs operativos.

Figura 19

Puntuación media en criterios técnicos por herramienta



Nota. Cada valor corresponde al promedio de las puntuaciones (escala 1–5) otorgadas por el supervisor en los cinco criterios técnicos para cada herramienta GenAI.

ChatGPT destaca con la mejor valoración técnica global (4,4), reflejando un equilibrio superior en relevancia, precisión, aplicabilidad, integridad y claridad estructural de sus recomendaciones. Le sigue de cerca Claude (4,2), mientras que Gemini y DeepSeek comparten una puntuación intermedia (3,4). Infinity, con 2,6, muestra el menor desempeño en este análisis técnico, indicando oportunidades de mejora en la calidad y profundidad de sus respuestas.

4.2 PROPUESTA DE LA HERRAMIENTA MÁS ADECUADA PARA LA GESTIÓN DE PROYECTOS

4.2.1 Herramienta recomendada

La selección de la herramienta más adecuada surge de un análisis integral que consideró tanto datos operativos cuantitativos como percepciones cualitativas de los usuarios clave. En el estudio comparativo previo, las variaciones semanales de los KPIs del supervisor —tiempo de supervisión, eficiencia de planificación, volumen y tiempo de revisión— no arrojaron un vencedor absoluto: cada GenAI mostró fortalezas puntuales, pero también limitaciones en

determinadas métricas. Por ejemplo, Gemini lideró en la reducción máxima del tiempo de supervisión, mientras que DeepSeek obtuvo el mayor incremento en eficiencia de planificación, lo que evidencia la ausencia de una única “mejor” solución bajo un criterio aislado.

Consciente de esta complementariedad, el presente trabajo se enfocó en identificar la herramienta que, de forma más recurrente y equilibrada, encabezara los gráficos de variación porcentual de los KPI más relevantes y acumulase el mayor número de mejoras significativas. Se ha dejado de lado el análisis de la herramienta con peor desempeño, pues el objetivo era determinar la opción óptima, no rankear a todas las participantes en un extremo bajo.

Bajo este enfoque, ChatGPT Pro (GPT-4o) se erige como la solución recomendada. En el plano cuantitativo, ChatGPT ocupó posiciones destacadas en múltiples indicadores clave: elevó la productividad general del equipo en un 10 % (equivalente a +4,92 imágenes/hora) y aumentó la velocidad de etiquetado en 12 % (+0,09 imágenes/minuto), mientras consiguió la mayor reducción de la tasa de error en etiquetado (-51 %). Al mismo tiempo, mantuvo un equilibrio notable en las métricas de supervisión, disminuyendo el tiempo total de supervisión en 29 % (-410 minutos) y mejorando la eficiencia de planificación en 42 % (+6,02 puntos porcentuales). Además, su latencia de respuesta promedio de 21 segundos (la más baja del estudio) permitió incorporar recomendaciones prácticamente al instante, sin interrumpir el flujo de trabajo en entornos de alta dinámica.

En cuanto a la valoración de usuarios, ChatGPT obtuvo la puntuación media global más alta tanto en la encuesta del supervisor (4,3/5) como en la del equipo (4,08/5), lo que refleja una percepción muy favorable de la claridad, relevancia y aplicabilidad de sus sugerencias. Del mismo modo, lideró la evaluación técnica en los cinco criterios definidos con una nota técnica promedio de 4,4/5, lo que refuerza la solidez y coherencia de sus recomendaciones más allá de los KPIs operativos.

Por todo ello, se propone ChatGPT Pro (GPT-4o) como la herramienta más adecuada para la gestión de proyectos en este caso de estudio. Su capacidad para generar insights rápidos, prácticas de optimización basadas en datos reales y recomendaciones de alta calidad facilita la toma de decisiones estratégicas y la mejora continua de procesos. En este escenario de recolección y etiquetado de imágenes, ChatGPT Pro demostró ser la opción más potente, ocupando las primeras posiciones en la mayoría de los KPIs clave y recibiendo las valoraciones más altas tanto del supervisor como del equipo. Gracias a su agilidad en la respuesta y a la profundidad de su análisis, es muy probable que esta solución también sobresalga en proyectos

de naturaleza similar —y potencialmente en otros contextos de gestión de proyectos— donde la rapidez y la solidez de los insights resulten determinantes para el éxito.

Más allá de los KPIs operativos, ChatGPT Pro destaca por su capacidad de comprender contextos complejos y generar respuestas altamente precisas y matizadas. A diferencia de otras GenAI que tienden a ofrecer recomendaciones más genéricas o enfocadas en un único aspecto, ChatGPT combina una amplia base de conocimiento con un modelo de lenguaje entrenado para mantener coherencia y detalle en explicaciones extensas. Esto le permite no solo presentar posibles mejoras, sino argumentar de manera lógica el porqué de cada sugerencia, anticipando contraargumentos y proponiendo alternativas. Su potencial radica en la flexibilidad para adaptar el tono y el nivel de profundidad según el interlocutor —desde resúmenes ejecutivos hasta planes de acción detallados—, así como en la capacidad de integrar información técnica y cualitativa en un solo output. Estas características hacen de ChatGPT Pro la herramienta con mayor potencial de precisión y utilidad estratégica frente a las demás opciones evaluadas.

Es fundamental subrayar que, aunque ChatGPT Pro ofrece un apoyo extraordinario en el análisis de datos, la generación de proyecciones y la formulación de planes de acción, no sustituye la figura del project manager. La gestión de proyectos implica competencias que trascienden el procesamiento de información: el liderazgo para motivar y coordinar equipos diversos, el juicio crítico para evaluar el contexto organizacional y adaptar soluciones a las políticas internas, y la capacidad de gestionar riesgos e imprevistos que surgen de manera ineludible en proyectos reales. El project manager aporta la visión estratégica de largo plazo, integrando factores humanos, técnicos y financieros en un todo coherente, y ejerce la autoridad de decisión última sobre las acciones recomendadas por la IA.

Además, el gestor asume la responsabilidad de validar la pertinencia de cada sugerencia, matizarla según las prioridades del negocio y garantizar su alineamiento con la cultura y normativas de la organización. La IA carece de la sensibilidad interpersonal necesaria para detectar tensiones de equipo, gestionar conflictos o fomentar un clima de colaboración, así como de la experiencia para anticipar riesgos reputacionales, regulatorios o de mercado que solo un profesional con trayectoria puede identificar. Por tanto, la integración de ChatGPT Pro debe verse como la incorporación de un copiloto cognitivo que potencia la eficiencia operativa —al liberar tiempo y reducir tareas mecánicas—, pero que siempre requiere del project manager para aportar contexto, juicio estratégico y habilidades de liderazgo, asegurando que la tecnología enriquezca el proceso de gestión sin desdibujar el papel insustituible del factor humano.

4.3 IMPACTO DE LA APLICACIÓN DE GENAI EN LA GESTIÓN DE PROYECTOS

4.3.1 Impacto en procesos y gestión

La adopción de ChatGPT Pro como soporte en el ciclo de vida del proyecto ha transformado sustancialmente la manera en que se diseñan, supervisan y ajustan las actividades operativas. Al automatizar el análisis de datos de rendimiento —desde tiempos de supervisión hasta métricas de calidad de etiquetado—, ChatGPT ha liberado al project manager de gran parte de las tareas de cálculo y consolidación de información, permitiéndole centrar sus esfuerzos en la planificación estratégica, la gestión de riesgos y la coordinación de equipos. Esta reducción de la carga administrativa intensiva promueve una gestión más proactiva: el gestor puede anticipar cuellos de botella, balancear recursos y realinear prioridades con mayor rapidez, en lugar de dedicarse a compilar estadísticas manualmente.

Además, ChatGPT Pro actúa como un copiloto cognitivo que sugiere mejoras prácticas para optimizar procesos y elevar la eficiencia del equipo. Gracias a su capacidad de integrar datos cuantitativos y cualitativos, la herramienta propone secuencias de trabajo que eliminan actividades redundantes, agrupan tareas similares y recomiendan patrones de loteo que mejoran el flujo operativo. Al generar planes de acción claros y priorizados —acompañados de la lógica detrás de cada propuesta—, ChatGPT facilita la comunicación entre el supervisor y los colaboradores, estableciendo un lenguaje común basado en métricas y objetivos. Este enfoque, centrado en la evidencia, refuerza la toma de decisiones colaborativa e incrementa la transparencia sobre los criterios que sustentan cada ajuste de proceso.

Por otra parte, la elaboración automatizada de reportes, y resúmenes ejecutivos acelera la validación de alternativas de gestión. El project manager puede evaluar distintas configuraciones de recursos o calendarios de trabajo en cuestión de segundos, comparando instantáneamente sus impactos potenciales sin necesidad de recurrir a herramientas de hoja de cálculo complejas. Este dinamismo en la simulación de escenarios fortalece la capacidad de respuesta ante imprevistos, ya que se reducen los tiempos de reacción y se aumenta la confianza en las decisiones adoptadas.

No obstante, es importante señalar que la eficacia de ChatGPT Pro radica en su uso como complemento del juicio humano. Aunque la IA sintetiza información y propone ajustes tácticos, el project manager conserva la responsabilidad de validar y contextualizar cada recomendación, adaptándola a las políticas internas, los acuerdos con los clientes y las particularidades del equipo. En este sentido, la combinación de la visión estratégica del gestor y el apoyo analítico de

ChatGPT crea un paradigma de gestión híbrido en el que se potencian tanto la precisión de los datos como la experiencia y el liderazgo humano.

En definitiva, la implementación de GenAI ha permitido optimizar los ciclos de planificación y control, agilizar la comunicación de resultados y fomentar una cultura de mejora continua basada en la medición objetiva de los procesos. Al actuar como un asistente inteligente, las herramientas de Inteligencia Artificial Generativa no solo aceleran las tareas operativas, sino que impulsan al project manager a enfocarse en las decisiones de mayor impacto, consolidando un modelo de gestión de proyectos más ágil, adaptable y fundamentado en evidencia.

4.4 RECOMENDACIONES PARA LA APLICACIÓN DE GENAI EN LA GESTIÓN DE PROYECTOS

4.4.1 Buenas prácticas para el uso de GenAI en proyectos

A partir de las lecciones aprendidas en este estudio, a continuación, se ofrecen pautas y factores clave para maximizar el valor de las herramientas de Inteligencia Artificial Generativa en la gestión de proyectos, garantizando su adopción efectiva y sostenible.

1. Definir objetivos claros y casos de uso: Antes de integrar GenAI, identifique las áreas del proyecto que más se benefician (planificación, control de calidad, etc.) y establezca indicadores de éxito para cada uno.
2. Estandarizar y asegurar la calidad de los datos: Es importante definir un conjunto de indicadores clave de su proyecto (tiempos, volúmenes, tasas de error, costos, productividad o cualquier otra métrica relevante) y alimente la IA desde un repositorio central organizado y validado. Esto garantiza que las recomendaciones se basen en información precisa, estructurada (KPIs históricos, flujos de trabajo, documentación técnica), coherente y libre de sesgos, aplicable a cualquier tipo de proyecto.
3. Diseñar prompts precisos y replicables: Utilice plantillas de prompts que incluyan contexto del proyecto, restricciones operativas y formato de salida esperado. Esto garantiza coherencia entre iteraciones y facilita la comparación de resultados.
4. Integrar validación humana continua: Establezca checkpoints en los que el project manager revise y ajuste las sugerencias de la IA, asegurando su pertinencia y alineación con políticas internas, cultura organizacional y prioridades estratégicas.
5. Formar y empoderar al equipo: Capacite tanto a project managers como a colaboradores en el uso de GenAI: redacción de prompts, interpretación de salidas

y prácticas de validación. Un equipo entrenado reduce la curva de adopción y mejora la calidad de los resultados.

6. Garantizar la seguridad y la confidencialidad: Defina políticas claras sobre datos sensibles y configure controles de acceso. Cuando trabaje con información propietaria, considere instancias on-premise o entornos seguros en la nube.

4.4.2 Factores críticos de éxito

Para asegurar que las recomendaciones generadas por GenAI se traduzcan en mejoras reales y sostenibles en la gestión de proyectos, es clave atender a los siguientes factores:

1. Compromiso y liderazgo del project manager: El gestor debe asumir un rol activo en la revisión, ajuste y priorización de las sugerencias de la IA, validando su viabilidad y alineándolas con los objetivos estratégicos y la cultura de la organización.
2. Integración en los procesos existentes: Incorporar las recomendaciones de GenAI en reuniones de planificación, revisiones diarias y rutinas de reporte, de modo que formen parte natural del ciclo de gestión y no queden aisladas.
3. Documentación y trazabilidad de decisiones: Registrar qué recomendaciones se implementaron, y cuál fue su impacto real, para construir un histórico de aprendizaje y facilitar futuras iteraciones.
4. Medición continua de resultados: No basta con aplicar la sugerencia; es necesario medir su efectividad a través de KPIs dedicados (por ejemplo, comparar tiempos, errores o productividad antes y después de la implementación) y ajustar el uso de la IA en función de esos hallazgos.
5. Cultura de mejora continua: Fomentar la mentalidad de iterar sobre las recomendaciones, adoptar cambios rápidos y compartir los aprendizajes, de modo que la organización integre la IA como un socio estratégico en la gestión de proyectos.
6. Comunicación constante y transparente: Establecer canales de comunicación claros, frecuentes y bidireccionales entre todos los involucrados para asegurar que las recomendaciones, ajustes y aprendizajes se compartan de manera oportuna, evitando malentendidos y favoreciendo la alineación de expectativas.

CONCLUSIONES Y RECOMENDACIONES

CONCLUSIONES FINALES

- Marco de selección riguroso de herramientas

La investigación implementó un marco de evaluación riguroso que combinó criterios de funcionalidad, accesibilidad y casos de uso real, lo que permitió identificar con precisión cinco herramientas de GenAI adecuadas para la gestión de proyectos. Este proceso de selección demostró su eficacia al filtrar soluciones que encajan con las necesidades de entornos de Sistemas de Información, asegurando que sólo se evaluaran plataformas viables y representativas del mercado.

- Validación empírica mediante métricas objetivas

Se cumplió un experimento controlado y el uso de KPIs operativos claros; los resultados demostraron diferencias cuantificables en productividad, tiempos de supervisión, calidad de etiquetas y latencia de respuesta. La metodología probó ser robusta y replicable, confirmando la validez de emplear métricas objetivas para comparar herramientas de GenAI en proyectos reales.

- ChatGPT Pro como solución óptima

Entre las cinco plataformas analizadas, ChatGPT Pro (GPT-4o) se consolidó como la opción más equilibrada y eficaz, al liderar de manera recurrente tanto los indicadores de productividad y calidad como las valoraciones de usuarios clave. Su combinación de velocidad de respuesta, profundidad analítica y aceptación por parte de supervisor y equipo justifica su adopción como estándar en proyectos con requisitos de agilidad y precisión.

- Modelo híbrido de gestión IA–humano

Más allá de los números, la integración de ChatGPT Pro potenció la agilidad y calidad de la gestión sin desdibujar el papel central del project manager. La IA liberó tiempo de tareas mecánicas, facilitó la toma de decisiones basadas en datos y mejoró la colaboración, pero siempre bajo el juicio crítico y liderazgo humano. Este enfoque híbrido —IA como copiloto, gestor como piloto— se presenta como la fórmula óptima para proyectos dinámicos y de alta complejidad.

- Guía filtrar soluciones para adopción de GenAI en Sistemas de Información

Este estudio cierra la brecha de evidencia práctica en Sistemas de Información al validar el uso de GenAI en proyectos reales. Ofrece un marco replicable que incluye selección estratégica de herramientas, medición de KPIs operativos y cualitativos, y buenas prácticas de implementación. Así, proporciona a las organizaciones un protocolo claro para incorporar IA en TI y sienta las bases para futuras aplicaciones en planificación financiera, gestión de productos y atención al cliente.

RECOMENDACIONES FINALES

- Adopción de un marco sistemático de evaluación de GenAI

Se recomienda establecer un protocolo documentado que integre criterios técnicos (capacidad de integración, rendimiento y escalabilidad), operativos (formatos de entrada/salida y latencia de respuesta) y de usuario (usabilidad y accesibilidad). Este marco debe probarse primero en proyectos piloto para validar su aplicabilidad antes de extenderlo a toda la organización.

- Definición de un modelo híbrido de colaboración IA–gestor de proyectos

Es imprescindible formalizar un proceso en el que la Inteligencia Artificial Generativa aporte análisis de datos, proyecciones y sugerencias de optimización, mientras que el project manager asuma la responsabilidad de validar, contextualizar y priorizar dichas recomendaciones dentro de las rutinas de planificación, seguimiento y control.

- Fortalecimiento de la gobernanza y calidad de los datos

Se recomienda centralizar toda la información del proyecto en un repositorio único, con estándares claros de nomenclatura, formatos y metadatos, así como controles de acceso y auditorías periódicas. Esta medida garantizará que la IA utilice datos precisos y libres de sesgos, elevando la fiabilidad de sus salidas.

- Implementación de ciclos de retroalimentación y mejora continua

Se aconseja diseñar y monitorear indicadores de adopción (porcentaje de prompts automatizados), desempeño (impacto sobre KPIs clave) y satisfacción de usuarios internos. Asimismo, deben programarse revisiones regulares para ajustar los “prompts”, documentar lecciones aprendidas y optimizar progresivamente los flujos de trabajo.

- Impulso a la capacitación y a la investigación futura en GenAI

Es esencial desarrollar un programa de formación continua que cubra la elaboración de “prompts”, la interpretación crítica de resultados y la validación de recomendaciones. Paralelamente, se anima a la comunidad académica y profesional a profundizar en estudios que amplíen este enfoque a distintos sectores, metodologías de gestión y plataformas tecnológicas (Enterprise Resource Planning – ERP, Business Intelligence – BI), incluyendo evaluaciones éticas y análisis del impacto a largo plazo de la GenAI en la gestión de proyectos.

BIBLIOGRAFÍA

- Wells, R. (2023). *8 AI tools every project manager needs in 2024*. Forbes. Retrieved from <https://www.forbes.com/sites/rachelwells/2023/11/22/8-ai-tools-every-project-manager-needs-in-2024/>
- HubSpot. (2024). *AI in project management: How artificial intelligence can improve project planning, tracking, and collaboration*. HubSpot Blog. Retrieved from <https://blog.hubspot.com/marketing/ai-project-management>
- Project-Management.com. (2024). *AI in project management: Uses, impacts, & 2024 trends*. from <https://project-management.com/ai-project-management/>
- IBM. (2024). What is artificial intelligence (AI)? IBM. <https://www.ibm.com/cloud/learn/what-is-artificial-intelligence>
- McKinsey & Company. (2023, June). The economic potential of generative AI: The next productivity frontier. <https://www.mckinsey.com/mgi/our-research/technology-and-innovation/the-economic-potential-of-generative-ai-the-next-productivity-frontier>
- Project Management Institute. (2017). A guide to the project management body of knowledge (PMBOK® Guide) (6th ed.). Project Management Institute. <https://www.pmi.org/pmbok-guide-standards/foundational/pmbok>
- Project Management Institute. (2005). Tendencias de gestión de proyectos y tecnología. <https://www.pmi.org/learning/library/es-tendencias-proyecto-carteragestion-tecnologia-7515>
- Rodríguez, M., & Bastidas, M. (2013). Análisis comparativo: Una propuesta didáctica. *Revista de Didáctica de la Lengua y la Literatura*, (25), 45–60. https://www.researchgate.net/publication/260777242_Analisis_comparativo_Una_propuesta_didactica
- Parra Mora, E. M., & Bolívar Cazañas, A. (2018). El impacto de la automatización en el mejoramiento de procesos. *Revista de Ingeniería y Tecnología*, 15(2), 55–65. https://www.researchgate.net/publication/324543601_El_impacto_de_la_automatizacion_en_el_mejoramiento_de_procesos

- Sembly.ai. (2024). *Using AI in project management: Key tools and benefits*. from <https://www.sembly.ai/blog/ai-in-project-management-key-tools-and-benefits/>
- Targett, D. (2023). *How AI is transforming project management*. TechTarget. Retrieved October 29, 2024, from <https://www.techtarget.com/searchenterpriseai/feature/How-AI-is-transforming-project-management>
- Kolbjørnsrud, V., & Thomas, R. J. (2023). *How AI will transform project management*. Harvard Business Review. Retrieved October 29, 2024, from <https://hbr.org/2023/02/how-ai-will-transform-project-management>
- Project Management Institute. (2025). *Generative AI overview for project managers* [Curso en línea]. <https://www.pmi.org/shop/tb/p/-elearning/generative-ai-overview-for-project-managers/el083>
- Anthropic. (2025). *Claude 4 Sonnet and Claude Code*. [Introducing Claude 4 \ Anthropic](https://www.anthropic.com/news/claude-4)
- DataCamp. (2025). *DeepSeek-V3-0324: A Guide With Demo Project*. <https://www.datacamp.com/tutorial/deepseek-v3-0324>
- Google DeepMind. (2025). *Gemini 2.5 Pro*. <https://deepmind.google/models/gemini/pro/>
- Hammoq. (2025). *Infinity AI*. <https://hammoq.com/>
- OpenAI. (2024). *Hello GPT-4o*. <https://openai.com/index/hello-gpt-4o/>
- OpenAI. (2024). *GPT-4 Technical Report*. OpenAI. <https://openai.com/research/gpt-4>
- Lifewire (2024). *Who Owns ChatGPT?* <https://www.lifewire.com/who-owns-chatgpt-8659181>
- Anthropic (2025). *Claude 4 Release*. <https://www.anthropic.com/news/claude-4>
- AboutAmazon (2025). *Claude on Amazon Bedrock*. <https://www.aboutamazon.com/news/aws/anthropic-claude-4-opus-sonnet-amazon-bedrock>

- TechTarget (2025). What is Gemini 2.5 Pro?
<https://www.techtarget.com/whatis/feature/Google-Gemini-25-Pro-explained>
- DeepSeek API Docs (2025). <https://api-docs.deepseek.com/news/news250325>
- DeepSeek-VL Team. (2024). *DeepSeek-VL: A Vision-Language Model with Reasoning Capabilities*. arXiv. <https://arxiv.org/abs/2401.10020>
- TechCrunch (2025). Startups building GenAI for productivity.
<https://techcrunch.com/tag/genai>

ANEXOS

ANEXO A – Nota metodológica sobre el uso de herramientas de apoyo

- Durante la elaboración de este trabajo, se utilizaron herramientas de inteligencia artificial generativa como apoyo en la redacción, estructuración y revisión de contenidos. Su uso fue supervisado críticamente por el autor, sin comprometer la originalidad ni la integridad académica del estudio.