



UNIVERSIDAD CÉSAR VALLEJO

ESCUELA DE POSGRADO

**PROGRAMA ACADÉMICO DE MAESTRÍA EN
INGENIERÍA DE SISTEMAS CON MENCIÓN EN
TECNOLOGÍAS DE LA INFORMACIÓN**

**Uso de chatGPT en la calidad del desarrollo web de una empresa
privada de Lima, 2023**

TESIS PARA OBTENER EL GRADO ACADÉMICO DE:

Maestro en Ingeniería de Sistemas con Mención en tecnologías de la
Información

AUTOR:

Robles Lujan, Darwin Edwin (ORCID.ORG/ 0000-0002-3306-2270)

ASESOR:

Mtr. Alza Salvatierra, Silvia Del Pilar (orcid.org/0000-0002-7075-6167)

Dr. Vargas Huaman, Jhonatan Isaac (orcid.org/0000-0002-1433-7494)

LÍNEA DE INVESTIGACIÓN:

Sistema de Información y Comunicaciones

LÍNEA DE RESPONSABILIDAD SOCIAL UNIVERSITARIA:

Desarrollo económico, empleo y emprendimiento

LIMA - PERÚ

2024

Dedicatoria

El presente trabajo dedico a mi familia quienes contribuyeron con sus consejos para culminar con la investigación, pese a las adversidades durante el tiempo de la maestría.

Agradecimiento

Agradezco a mi maestra Silvia Alza por la paciencia y la enseñanza meticulosa durante el desarrollo de mi investigación. Pese a los contratiempos supo entender y ayudarme en el paso a paso de la investigación. Sin ella no hubiera sido posible este estudio.

Índice de contenidos

Dedicatoria	ii
Agradecimiento	iii
Índice de contenidos	iv
Resumen	v
Abstract	vi
I. INTRODUCCIÓN	1
II. MARCO TEÓRICO	4
III. METODOLOGÍA	13
3.1 Tipo y diseño de investigación	13
3.2 Variables y operacionalización	14
3.3 Población, muestra y muestreo	16
3.4 Técnicas e instrumentos de recolección de datos	17
3.5. Procedimientos	18
3.6 Método de análisis de datos	20
3.7 Aspectos éticos	21
IV. RESULTADOS	22
V. DISCUSIÓN	25
VI. CONCLUSIONES	31
VII. RECOMENDACIONES	32
REFERENCIAS	33
ANEXOS	40

Resumen

El propósito fundamental de este trabajo consiste en examinar los efectos del empleo de ChatGPT en la calidad del desarrollo web en una empresa privada de Lima durante el año 2023. La problemática abordada se enfoca en la incertidumbre acerca de cómo la incorporación de esta herramienta basada en inteligencia artificial puede incidir en la eficiencia, seguridad y rendimiento del proceso de desarrollo de aplicaciones web. La metodología utilizada para abordar esta investigación se cataloga como aplicada y se sustenta en un diseño cuasiexperimental. Durante la creación de una aplicación web, se recopiló datos comparando dos grupos: uno que empleó ChatGPT y otro que prescindió de su uso. El estudio adoptó un enfoque cuantitativo, aplicando la técnica de fichaje para visualizar los resultados consignados en fichas de control. Se aplicaron herramientas de prueba de calidad de código, como SonarQube y Sonar Scanner, para evaluar elementos como vulnerabilidades, Code Smells, Bugs, código duplicado y Security Hotspots que inciden en la calidad del desarrollo web.

En base a los resultados obtenidos, se determinó un incremento de la calidad del desarrollo web bajo un porcentaje del 99.12%, en lo que respecta a escalabilidad se obtuvo un aumento del 434.13%, en base a la seguridad se obtuvo un aumento del 80% y para el rendimiento se determinó un aumento 1.37% llegando a la conclusión de que el uso del ChatGPT aumenta la calidad del desarrollo web en una empresa privada en Lima 2023.

Palabras clave: ChatGPT, sistema web, calidad de desarrollo Web, sistemas de información, SonarQube

Abstract

The main purpose of this work is to examine the effects of the use of ChatGPT on the quality of web development in a private company in Lima during the year 2023. The problem addressed focuses on the uncertainty about how the incorporation of this tool based on artificial intelligence can affect the efficiency, security and performance of the web application development process. The methodology used to approach this research is classified as applied and is based on a quasi-experimental design. During the creation of a web application, data were collected comparing two groups: one that used ChatGPT and another that did not use it. The study adopted a quantitative approach, applying the clocking technique to visualize the results recorded on control cards. Code quality testing tools, such as SonarQube and Sonar Scanner, were applied to evaluate elements such as vulnerabilities, Code Smells, Bugs, duplicate code and Security Hotspots that affect the quality of web development.

Based on the results obtained, it was determined an increase in the quality of web development under a percentage of 99.12%, regarding scalability an increase of 434.13% was obtained, based on security an increase of 80% was obtained and for performance an increase of 1.37% was determined, reaching the conclusion that the use of ChatGPT increases the quality of web development in a private company in Lima 2023.

Keywords: ChatGPT, web system, web development quality, information systems, SonarQube

I. INTRODUCCIÓN

En el marco internacional la evolución de la tecnología del siglo XXI ha llevado a la aparición de nuevas herramientas y recursos que facilitan el desarrollo de aplicaciones. En el entorno de calidad de desarrollo web el advenimiento de las nuevas tecnologías facilita que los hackers puedan vulnerar las aplicaciones web. En el ámbito de la programación el uso de ChatGPT, una herramienta basada en inteligencia artificial y lenguaje natural, cambió la forma de codificar un lenguaje de programación desde inicios del mes de diciembre del 2022. Esta investigación tiene como objetivo, estudiar el impacto del uso de ChatGPT en la calidad de desarrollo de un aplicativo web en la empresa Investigación S.A.C 2023.

La realidad problemática internacional de la investigación sobre la influencia del uso de ChatGPT según Cooper, G. (2023), radica que ChatGPT corre el riesgo de posicionarse como la máxima autoridad epistémica, donde se asuma como una única verdad, sin fundamentación adecuada en pruebas ni presentación con suficientes calificaciones, Que podría afectar la calidad de la codificación de software y aplicaciones web.

Por su parte Bozkurt, A. (2023), indica que es probable que la IA generativa catalice la convergencia de numerosas tecnologías convencionales y educativas, y requiere nuestra preparación para un futuro dominado por la IA. Hay casos en los que la inteligencia artificial generativa ha sido criticada por generar información superficial o falsa y no estar a la altura de las expectativas iniciales. Sin embargo, es importante reconocer que esta limitación no es exclusiva de la IA, ya que los seres humanos también se les consideran falibles e imperfectos. Ante esta realidad los programadores de Investigación S.A.C no son excluidos del uso de la herramienta chatGPT para resolver las tareas día a día y mejorar la calidad de desarrollo web.

En el contexto local, el proceso de codificación y las habilidades del programador son fundamentales para los ingenieros de Investigación S.A.C. Tradicionalmente, los programadores durante el desarrollo han recurrido a una combinación de recursos como tutoriales, manuales, ejemplos en Stack Overflow durante el desarrollo de un Software con el objetivo de mejorar la calidad del

desarrollo. Sin embargo, la llegada de herramientas basadas en inteligencia artificial, como ChatGPT, plantea nuevas posibilidades en cuanto a la forma en que los programadores pueden interactuar y desarrollar un software. Por ejemplo, los programadores junior naturalmente acceden a ChatGPT para consultar cualquier duda en la lógica de un algoritmo.

A medida que los programadores se enfrentan a desafíos y preguntas durante el proceso de desarrollo de una web, el uso de ChatGPT puede ofrecer una nueva forma de obtener respuestas y orientación inmediata de soluciones de código. Al proporcionar una interacción conversacional similar a la que tendrían con un tutor o compañero con experiencia, esta herramienta tiene vigor para optimizar la eficiencia y la calidad del aprendizaje de la programación, según Yilmaz, R., & Yilmaz, F. G. K. (2023)

Sin embargo, es importante investigar y evaluar la influencia real del uso de ChatGPT en la calidad del desarrollo del aplicativo. Es importante analizar cómo esta tecnología afecta la calidad de desarrollo de un aplicativo web en la empresa Investigación S.A.C 2023.

La presente investigación se enmarca en el estudio sobre el impacto del uso de la herramienta ChatGPT en el desarrollo de un aplicativo web de Investigación S.A.C. Para comprender mejor este tema, se exploraron diversas fuentes que abordan conceptos clave relacionados con el uso de herramientas basadas en inteligencia artificial.

Según lo expuesto se plantea el siguiente problema general: ¿Cómo influye el uso de ChatGPT en la calidad de desarrollo de una web en la empresa Investigación S.A.C en 2023? Y como consiguiente los siguientes problemas específicos: (I) ¿Cuál es el efecto del uso de la herramienta de ChatGPT en la escalabilidad de desarrollo de un aplicativo web en la empresa Investigación S.A.C 2023?, (II) ¿Cómo influye el uso de la herramienta de ChatGPT en la seguridad del desarrollo de un aplicativo web en la empresa Investigación S.A.C 2023?, (III) ¿Cuál es el impacto del uso de ChatGPT en el rendimiento del desarrollo de un aplicativo web en la empresa Investigación S.A.C 2023?

El objetivo general de la investigación es determinar la influencia del uso de ChatGPT en la calidad de desarrollo web en la empresa Investigación S.A.C en 2023. Por tanto, se plantea los siguientes objetivos específicos: (a) determinar

el efecto del uso de ChatGPT en la escalabilidad de una web en la empresa Investigación S.A.C 2023, (b) analizar la influencia del uso de ChatGPT en la seguridad en el desarrollo de una web en la empresa Investigación S.A.C 2023, (c) determinar el impacto del uso de ChatGPT en el rendimiento de la aplicación web en la empresa Investigación S.A.C de Lima Norte en 2023.

La hipótesis general es: que el uso de ChatGPT tendrá un impacto positivo en la calidad del desarrollo web en la empresa Investigación S.A.C, mejorando la eficiencia, la precisión y la velocidad en la creación y optimización de códigos. Y las hipótesis específicas son: (A) el uso de ChatGPT en el desarrollo de una web en la empresa Investigación S.A.C 2023 tendrá un efecto positivo en la escalabilidad de la web, (b) el uso de ChatGPT durante el desarrollo de una web en la empresa Investigación S.A.C 2023 influirá positivamente en la seguridad de la web y (c) el uso de ChatGPT en el desarrollo web de la empresa Investigación S.A.C de Lima Norte en 2023 tendrá un impacto positivo en el rendimiento de la web.

El estudio se justifica porque aporta información para que las empresas impulsen el uso de chatGPT en el desarrollo web cuidando las desventajas y ventajas y además sirve a la comunidad académica para que tome dentro de la formación para los programadores las buenas prácticas de codear con la ayuda de chatGPT.

II. MARCO TEÓRICO

Sobre el uso de chatGPT en el desarrollo web los investigadores que antecedieron al este estudio llegaron a las siguientes conclusiones:

Según la investigación de los autores Ozturk et al. (2023) ChatGPT tiene el 91 % de falsos positivos en el análisis de las vulnerabilidades del código estático en comparación a cualquier analizador de código estático tradicional que es del 82%. Estos hallazgos resaltan el prometedor potencial de ChatGPT para mejorar el proceso de análisis de código estático, pero revelan ciertas advertencias en la precisión en su estado actual. Los hallazgos del estudio sugieren que una posibilidad interesante para explorar en trabajos futuros sería elegir lo mejor de ambos mundos, combinando analizadores de código estático tradicionales con ChatGPT para encontrar vulnerabilidades de seguridad de manera más efectiva.

Por su parte Tian et al. (2023) concluyen que ChatGPT tiene una capacidad de atención limitada en la reparación de tareas basadas en la semántica artística. En tanto en las tareas de resumen de código puede producir explicaciones de intención coherentes para códigos incorrectos frente a la verdad básica y que ChatGPT puede razonar sobre el original con mejor capacidad. Esta información brinda la oportunidad de aprovechar ChatGPT en el campo de investigación en el entorno de prueba.

Otro estudio de Yilmaz & Yilmaz (2023) indica que los estudiantes afirmaron que el uso de ChatGPT en el proceso de programación mejoró efectivamente sus habilidades de pensamiento. Significa que el uso de chatGPT influye en los resultados de desarrollo de una aplicación y mejora del pensamiento lógico del usuario (programador).

Se tiene en cuenta otra investigación interesante de Fajkovic, & Rundberg (2023) que concluye tango Copilot y ChatGPT entregaron código de calidad similar, ambas herramientas lograron crear los sitios web de acuerdo con una estructura alámbrica con pequeñas diferencias de estilo. Descubrieron que es más fácil crear un sitio web desde cero utilizando ChatGPT que con Copilot de GitHub. Sin embargo, la IA generadora de código no está lo suficientemente avanzada como para crear sistemas desde cero de manera eficiente y sin introducir errores ni riesgos de seguridad.

ChatGPT como herramienta de apoyo: ChatGPT es un instrumento de IA que permite a los usuarios interactuar con un asistente virtual de texto. Según Savelka, Agarwal, Bogart, & Sakr (2023), puede generar código de programa de computadora, es decir, realizar síntesis de programas informáticos con un alto grado de éxito. También pueden explicar el programa de computadora código en términos de lenguaje natural.

Los autores Al-Khiami & Jaeger (2023) en su investigación experimental concluyen que pedir el código completo de aplicación desde principio al chatGPT terminó con códigos inconclusos, sin embargo, cuando solicitaban funciones específicas entrega un código correcto, por tanto, el desarrollo con chatGPT debe ser por fases.

Por su parte el investigador York, E. (2023) tras resaltar que chatGPT es un buen asistente para un junior, cuestiona que es necesario centrar más la investigación en su capacidad organizativa y evaluativa del chatGPT porque falta validar la capacidad organizativa y priorización de tareas y subtareas en UX.

En tanto los autores White et al. (2023) describen en su investigación que LLMs como chatGPT tienen un inmenso potencial para ayudar a automatizar tareas comunes a lo largo de la secuencia de fases en la práctica de la ingeniería de software, potencial para acelerar y no sólo generando código, sino haciendo experimentaciones rápidas en muchos niveles. Puntualizan que es clave aprovechar las capacidades de codificar con un catálogo eficaz de indicaciones y orientación sobre cómo combinarlos en diferentes etapas del proceso de desarrollo de software para mejorar la ingeniería. Otros investigadores Backström & Kihlert (2023) según su estudio indican una mejora en la calidad general del código cuando se utilizó ChatGPT para mejorar el código escrito por los estudiantes. Detalla que la medida en que ChatGPT pudo eliminar las violaciones de calidad del código es difícil de determinar. Sin embargo, el estudio identificó que ChatGPT logró identificar las vulnerabilidades y mejorar la calidad del código. ChatGPT tenía más probabilidades de abordar y resolver ciertas infracciones que otras. Esto implica que la medida en que ChatGPT puede mejorar la calidad del código ante las infracciones también dependen del tipo de infracciones que se estudian.

El autor Lesbekkyzy (2023) en una de sus investigaciones indica que las pruebas de software en ChatGPT es eficiente y la ejecución es efectiva en casos de prueba (test-QA), porque reduce el tiempo y el costo de las pruebas, y mejora la confiabilidad y repetibilidad del proceso de prueba.

El autor Biswas (2023) en un artículo publicado concluye que ChatGPT es una herramienta poderosa para la comunidad de programadores, posee el potencial para generar un impacto sustancial en el campo porque puede mejorar la eficiencia y la precisión en la programación.

La investigación de los autores Liu et al. (2023) visibiliza que, si bien ChatGPT puede generar código funcional para diversas tareas de programación, el código generado a menudo sufre de problemas de calidad, como compilación y errores, resultados incorrectos, problemas de mantenimiento y ineficiencias en el desempeño.

Por su parte los investigadores Feng et al. (2023) revelan en un estudio de análisis de sentimientos que las personas generalmente tienen temores sobre las capacidades de generación de código de ChatGPT, en lugar de sentirse feliz, enojado, sorprendido o triste.

En un estudio los investigadores Yetiştirten et al. (2023) realizaron una comparación de la efectividad en la generación de código de GitHub Copilot, Amazon CodeWhisperer y ChatGPT, donde chatGPT tuvo la tasa de éxito más alta entre los evaluados, fue capaz de generar soluciones de código correctas para 65,2 % de los problemas en el conjunto de datos de problemas de HumanEval. También produjo parcialmente soluciones correctas para el 22,6% de los problemas y soluciones incorrectas para el 12,2% de los problemas, según los investigadores.

En tanto los investigadores Lau & Guo (2023) tras realizar un estudio dejan las siguientes inquietudes sobre si las herramientas de codificación de IA se vuelven más omnipresentes en el futuro, ¿qué los temas pedagógicos eternos deberían seguir siendo los mismos?, ¿Cómo se podrán preparar a nuestros estudiantes para los próximos 50 años? Y ¿Qué necesitarán saber los programadores en el año 2073? La última pregunta aún no tiene una hipótesis clara pese a las nuevas funcionalidades que sacó OpenAI en su última presentación del mes de noviembre del 2023.

En el proceso de aprendizaje de programadores los autores Waseem et al. (2023) según sus hallazgos enfatizan la necesidad de un enfoque equilibrado al integrar instrumentos impulsados por la inteligencia artificial como ChatGPT en el desarrollo de software. Mientras ofrecen beneficios sustanciales, los desarrolladores deben mantener una mentalidad independiente para la resolución de problemas (parte fundamental para mejorar la calidad de desarrollo).

El investigador Tasic (2023) muestra la preocupación de la dependencia de las empresas y las personas a las herramientas impulsadas por la IA, lo que podría conducir a una pérdida de ciertas habilidades y toque humano. Además, los modelos de IA pueden perpetuar involuntariamente y amplificar los sesgos existentes encontrados en sus datos de entrenamiento. Esto podría llevar al desarrollo de aplicaciones web sesgadas o injustas, lo que podría conducir a la exclusión o la discriminación, aparte de no contribuir en el cuidado del medio ambiente por recursos energéticos que necesita.

El estudio de los autores Szabó & Bilicki (2023) sobre el enfoque de seguridad de aplicaciones web muestra que GPT-4 obtuvo una precisión del 88,76%, en el análisis de código estático y las capacidades de interpretación de modelos de lenguaje grandes durante la investigación que desarrollaron los autores.

En un estudio comparativo de generación de código los autores Hansson & Ellréus (2023) revelan que chatGPT como Copilot suelen generar código sin errores de calidad. ChatGPT según el estudio generó el 80,7% de los algoritmos sin ninguna infracción de las reglas de calidad, mientras Copilot generó 64,7 % de algoritmos sin vulnerabilidad. Por lo tanto, tanto ChatGPT como Copilot parecen funcionar bien en la generación de código de alta calidad. Sin embargo, chatGPT parece funcionar mejor que Copilot y el análisis mostró una diferencia estadísticamente significativa entre las dos herramientas, según los investigadores.

Por su lado el investigador Hörnemalm (2023) alerta a través de los resultados de su estudio que los desarrolladores junior confían en ChatGPT más que en los desarrolladores senior, lo que muestra una correlación entre el nivel de experiencia y el riesgo potencial al usar IA generativa. Por lo tanto, es

fundamental que las empresas que opten por utilizar chatGPT necesiten un sistema de soporte sólido por los riesgos.

Amri (2023) según su investigación concluye que la integración de la IA en el desarrollo de sitios web ha transformado la industria al mejorar la experiencia del usuario, permitiendo análisis de datos inteligentes, automatización de la creación de contenido, mejora del diseño web, fortalecimiento de la seguridad y optimización SEO. A medida que la IA siga evolucionando, dará forma al futuro del desarrollo de sitios web, según la investigadora. Como se pudo visualizar en todos los antecedentes aún existen discusiones entre la calidad y los riesgos de la generación de código con chatGPT.

Según los investigadores Andersson & Marshall (2023) ChatGPT podría ayudar a los desarrolladores con la depuración de código, generación de código, refactorización, optimización, educación de los desarrolladores junior, documentación técnica y actuar como socio colaborador durante la lluvia de ideas. Sin embargo, ChatGPT puede tener limitaciones potenciales cuando se trata de código complejo. Por tanto, incluyen preocupaciones sobre la integridad y la seguridad, particularmente con el código confidencial, la falta de conocimiento sobre las funciones de ChatGPT y las limitaciones funcionales, según las conclusiones de los investigadores.

En tanto la gestión de proyectos de software, los autores Abbas et al.(2023) visibilizan que la aplicación de chatGPT en la dirección de proyectos de software aporta un conjunto de ventajas que deben equilibrarse con sus posibles limitaciones. Si bien chatGPT puede mejorar significativamente la comunicación, automatizar tareas rutinarias y proporcionar información valiosa, se debe reconocer que carece de inteligencia emocional y puede generar información incorrecta o engañosa si no se guía correctamente. También se debe tener en cuenta las consideraciones éticas sobre la privacidad y la necesidad de experiencia humana en determinados ámbitos.

Existe otra preocupación de seguridad y evidenciar el problema los autores Pa Pa et al. (2023) alertan que ChatGPT, Auto-GPT y text-davinci-003, demuestran el potencial para generar malware y herramientas de ataque bajo control de seguridad y moderación, lo que destaca la importancia de mejorar las medidas de seguridad y los controles de seguridad en sistemas de IA.

Los resultados del estudio de los autores Zhang et al. (2023) sobre la reparación automatizada de programas demuestran que chatGPT genera 109 parches correctos sobre 151 errores cuando solo se le da el mensaje básico. Además, chatGPT continúa corrigiendo 18, 25 y 10 errores adicionales con mensajes que contienen descripciones de problemas de programación, mensajes de error y localización de errores. Al participar en diálogos, chatGPT soluciona nueve errores más. Estos resultados indican que ChatGPT tiene una capacidad prometedora para corregir errores, que puede mejorarse aún más con indicaciones adecuadas y más diálogos.

Por otro lado, un estudio de los investigadores Xie et al. (2023) que un test unitario realizado con chatGPT supera a EvoSuite en términos de cobertura de líneas y sucursales y supera a AthenaTest y A3Test en cobertura de métodos focales, significa que las pruebas de QA podrían ser mejores con la automatización de los tests unitarios con chatGPT.

Otros autores como Sakib et al. (2023) revelan que si bien el chatGPT destaca en la generación de soluciones, no produce constantemente las soluciones más eficientes en términos de tiempo de ejecución o uso de memoria. Aproximadamente el 30% de las veces genera soluciones inexactas que no se pueden rectificar ni siquiera con los comentarios proporcionados por Leetcode. Esto resalta las limitaciones del modelo en ciertos escenarios problemáticos. Sin embargo, esta exploración de las capacidades de ChatGPT subraya su potencial para revolucionar la generación de código y ayudar a los programadores en sus tareas. Sirve como testimonio del poder de las herramientas basadas en IA para aumentar los flujos de trabajo de codificación.

En un estudio sobre la adopción del uso de chatGPT por los programadores los investigadores Ge & Wu (2023) explican que chatGPT tiene el potencial de ayudar a los desarrolladores de software profesionales a corregir errores, pero aún no se ha adoptado ampliamente. Los hallazgos ofrecen información importante sobre los aspectos que inciden en la adopción de ChatGPT para corregir errores.

En el ámbito nacional, aún no existen investigaciones sobre el uso del chatGPT en el desarrollo de websites. Por tal motivo esta investigación presentará una base exploratoria para los futuros investigadores nacionales.

Se espera que esta revisión literaria contribuya a proporcionar un marco conceptual sólido para el estudio y permita identificar las lagunas en la literatura existente. Además, se utilizaron los conocimientos adquiridos para diseñar el enfoque metodológico de la investigación, seleccionar las variables relevantes y desarrollar los instrumentos de recolección de datos adecuados.

Los investigadores Bale, Vada, Oshiojum, Lakkineni, Rao, Venkatesh & Rani (2023) concluyen que chatGPT es útil para el desarrollo de código, la reparación de errores y las pruebas de software. También podría ayudar a los desarrolladores con actividades de procesamiento (ejecución) del lenguaje natural, como el análisis de requisitos y el diseño del espacio de interacción con usuario. Sin embargo, podría plantear dificultades éticas, como el riesgo de sesgo de datos, la divulgación de información privada y la puesta en peligro de datos privados. Los autores proponen utilizar ChatGPT para optimizar la interfaz de usuario y la comunicación de la interacción del usuario.

Calidad de desarrollo web: el desarrollo web se refiere las fases de creación y mantenimiento de sitios web y aplicaciones web. En un entorno empresarial, la calidad del desarrollo web es un aspecto fundamental, un sitio web actualmente es la primera impresión de los clientes sobre una empresa. La calidad se relaciona con la eficiencia, el rendimiento, la seguridad y la satisfacción del usuario. Según Smith & Sidaros (2018), la calidad del desarrollo web se puede medir a través de métricas objetivas, como el tiempo de carga de una página, la disponibilidad y la resistencia a ataques de seguridad. También se puede evaluar subjetivamente a través de la usabilidad y la satisfacción del usuario.

La teoría de la Inteligencia Artificial General conocida como IAG según los autores Wang & Goertzel (Eds.). (2012) propone que la IA debería aspirar a crear sistemas que puedan realizar cualquier actividad intelectual que un ser humano pueda y sobre el caso los autores Latapie et al. (2021) acuñan que IAG puede arrojar luz sobre el camino hacia una mejor comprensión de cómo mejorar la cognición humana.

En tanto los autores Latapie et al. (2021) describen que la IAG implica la capacidad de lograr una variedad de objetivos y ejecutar tareas en contexto y entornos diferentes. Es un sistema con capacidad para resolver problemas muy diferentes previstos por sus creadores.

La segunda teoría es sobre la Red Neuronal profunda donde Elbrächter et al. (2019) exponen el uso de redes neuronales profundas para el aprendizaje automático y muestran la aproximación de funciones suficientemente suaves, las redes profundas de ancho finito requieren una conectividad estrictamente menor que las redes de ancho finito.

La tercera teoría es la racionalidad limitada donde Schwarz et al. (2022) discuten sobre la contribución de H. Simon quien desarrolló que los agentes racionales (incluidas las IA) toman decisiones basadas en recursos limitados, en lugar de maximizar completamente la utilidad.

Por otra parte, los resultados de la investigación de Petracca (2021) concluye que las representaciones (y los cálculos) de la racionalidad limitada propone retenerlas reformando su naturaleza abstracta, mientras que la racionalidad corporal y la racionalidad extendida se basan en formas atenuadas o intermitentes (es decir, dependiendo de la tarea cognitiva) de representacionalismo.

La variable dependiente denominada uso de ChatGPT, se define según Rahmaniar (2023), como la integración y aplicación de ChatGPT en el proceso de desarrollo web. Esta incorporación implica aprovechar ChatGPT como una herramienta valiosa que va desde sugerir fragmentos de código relevantes según la tarea actual hasta proporcionar asistencia en la depuración, actuando esencialmente como un programador asociado virtual. En términos simples, el "Uso de ChatGPT" se refiere a la utilización de esta herramienta para la generación de código durante el desarrollo web en tiempo real.

Una de las principales virtudes de los modelos GPT radica en su capacidad para generar un discurso coherente, fluido y prácticamente indistinguible del texto creado por seres humanos. Estos modelos han sido eficazmente empleados en diversas aplicaciones, tales como la creación de chatbots, generación de contenidos y traducción automática. Su destreza se manifiesta especialmente en la elaboración de respuestas a preguntas abiertas, consolidándose, así como una herramienta fundamental para la comunicación en lenguaje natural. La estructura neuronal basada en transformadores de GPT, a través de sus capas, utiliza técnicas de atención para focalizarse en áreas específicas del texto de entrada. Esta arquitectura permite al modelo captar

patrones estadísticos inherentes al lenguaje natural, prescindiendo de la necesidad de programar de manera explícita reglas sintácticas o gramaticales (Abdullahi et al, 2024).

Por tanto, la variable independiente de la investigación es la calidad del desarrollo web, hace referencia a la excelencia, eficiencia y conformidad con estándares durante el proceso de creación de un sitio web. Esta dimensión abarca diversos aspectos, tales como rendimiento, seguridad, escalabilidad y la ausencia de errores o inconvenientes técnicos, según lo definido por Pargaonkar, S (2022).

La calidad del software es un tema que requiere mucha atención y dedicación. Sin embargo, es raro encontrar software que sea perfecto. Cada proyecto se esfuerza por crear software de calidad excepcional que no sólo cumpla, sino que también supere las expectativas del usuario, si es posible (Marcillo y Román, 2022). Por otro lado, la escalabilidad, está intrincadamente entrelazada con el diseño de un sistema. Tiene un profundo impacto en el rendimiento, cuando una aplicación está bien diseñada, no hay necesidad de preocuparse por problemas de escalabilidad. Se puede discernir la escalabilidad examinando cuidadosamente la implementación y el diseño general del sistema, pero, no es algo que pueda configurarse como un atributo del sistema (Sánchez y Ramírez, 2022). Asimismo, la seguridad del software implica la incorporación de medidas de protección durante la fase de desarrollo, con el objetivo de asegurar su resiliencia frente a posibles amenazas. En esencia, esto implica someter un software a rigurosas evaluaciones de seguridad antes de su lanzamiento, para evaluar su capacidad de resistir infiltraciones malévolas (Capuñay y Antón, 2021). No obstante, la prueba de software es el proceso de utilizar herramientas para garantizar que un sistema funcione correctamente. Este método también implica el uso de recursos como pruebas de rendimiento, que evalúan qué tan bien funciona un programa o aplicación bajo cargas de trabajo específicas, garantizando así la calidad del software (Veloz et al., 2021).

III. METODOLOGÍA

3.1 Tipo y diseño de investigación

3.1.1. Tipo de Investigación

Los autores Mishra & Alok (2022) definen que la investigación aplicada consiste en hallar respuestas concretas y práctica para abordar desafíos específicos que enfrenta un individuo, sociedad o una organización industrial o empresarial, por ejemplo, cómo para abolir los delitos de odio, cuáles son las formas de comercializar un producto, qué está provocando un aumento de la pobreza, etcétera. Por tanto, esta investigación se toma como aplicada porque busca una solución práctica.

3.1.2. Diseño de Investigación

Según los actores Miller et al. (2020) los diseños cuasiexperimentales permiten a los científicos realizar estudios rigurosos en diferentes contextos. Teóricamente, ambos grupos están expuestos a las mismas tendencias en el entorno, lo que hace posible descifrar si la intervención tuvo algún efecto. La medición tanto de las condiciones de tratamiento como de control ocurre clásicamente antes y después de la intervención, con una mejora diferencial entre los grupos atribuidos a la experimentación.

Este diseño es popular debido a su practicidad, especialmente si los puntos de recolección de datos se pueden mantener al mínimo. Puede ser especialmente útil para capitalizar experimentos como los que pueden ocurrir en el contexto de ciertas iniciativas o implementaciones de políticas, específicamente, implementaciones en las que es plausible que se pueda identificar un grupo de control.

El diseño de este estudio es cuasiexperimental, porque se han recopilado datos durante el proceso de creación de una aplicación web usando chatGPT y sin usar dicha herramienta para el proceso de codificación, es decir, que existe un grupo control y experimental. El objetivo fue medir la influencia de chatGPT en la calidad del desarrollo de una página web. El método que se utilizó fue el experimento controlado. Donde el grupo de control pertenece al desarrollo de web

sin el uso de chatGPT y el grupo experimental el desarrollo con el uso de chatGPT.

El enfoque del estudio fue cuantitativo, porque visibiliza de manera estadística la influencia de la variable independiente con la dependiente. Según los autores Mehrad & Zangeneh (2019) El estudio cuantitativo es aquella en la que el investigador utiliza principalmente afirmaciones postpositivistas para la evolución del conocimiento, por ejemplo; pensamiento de causa y efecto, reducción a variables específicas, hipótesis y preguntas, uso de mediciones y observaciones, y prueba de teorías. Las estrategias utilizadas frecuentemente en este diseño de investigación son experimentos y encuestas, e instrumentos predeterminados en la recopilación de datos que producen datos estadísticos.

3.2 Variables y operacionalización

Variable independiente: Uso de ChatGPT

Definición conceptual: Rahmaniar (2023) define a chatGPT como una gran herramienta para ayudar a los desarrolladores en tiempo real. Desde sugerir fragmentos de código relevantes basados en la tarea actual hasta asistencia de depuración, que puede actuar como un programador asociado virtual. El "Uso de ChatGPT" se refiere a la incorporación y aplicación en el proceso de desarrollo de una web. Implica la utilización de ChatGPT para generar código durante el desarrollo web.

Definición operacional: El uso de chatGPT se medirá mediante el registro del número de interacciones de chatGPT en el proceso de desarrollo web. Esto incluye la cantidad de contenido generado por ChatGPT, las sugerencias proporcionadas por la herramienta y la integración de funcionalidades basadas en ChatGPT en el sitio web.

Escala de medición: Para medir el uso de la herramienta chatGPT, se empleó una escala de medición que varía desde "No Uso" (0) hasta "Uso Extensivo" (1), donde 0 representa la ausencia de cualquier uso de chatGPT en el desarrollo web y 1 indica un uso en múltiples aspectos del proceso de desarrollo.

Variable dependiente: Calidad de desarrollo web

Definición conceptual: La calidad de desarrollo web se refiere a la excelencia, eficiencia y cumplimiento de estándares en la creación de una página web. Incluye dimensiones como rendimiento, seguridad, escalabilidad y la ausencia de errores o problemas técnicos según Pargaonkar (2022).

Definición operacional: La calidad de desarrollo web se medirá mediante una combinación de indicadores, que incluyen cantidad errores de la seguridad, la cantidad de errores de Code Smell, número de bugs identificados, número de vulnerabilidades identificados y número de código duplicado.

Escala de medición: Para medir la calidad de desarrollo web, se emplearon escalas de intervalo (razón) para cada indicador, como número de vulnerabilidades identificados, número de bugs identificados, número de security Hotspots, número de Code Smell, número de código duplicado, y una escala de puntuación de tiempo de respuesta del servidor, puntuación de rendimiento en Pagespeed, número de solicitudes de HTTP.

Variable independiente: Uso de ChatGPT

Según Brown et al. (2020) un modelo de lenguaje grande se entrena en un nivel suficientemente conjunto de datos grande y diverso, es capaz de funcionar bien en muchos dominios y conjuntos de datos. Se refiere a la implementación y aplicación de modelos de lenguaje basados en inteligencia artificial, como GPT-3 (Generative Pre-trained Transformer 3), desarrollados por OpenAI, para llevar a cabo interacciones de lenguaje natural y generar respuestas coherentes. Estas herramientas son capaces de comprender preguntas, solicitudes o comandos escritos por usuarios y proporcionar respuestas que simulan una conversación fluida con un interlocutor humano.

Variable dependiente: Calidad de desarrollo web

Pargaonkar (2022) define la calidad del código como fundamental y determinante en la confiabilidad (seguridad), escalabilidad y el rendimiento de un producto de software. Por tanto, la calidad es un conjunto de atributos y

estándares que determinan la excelencia y efectividad en la creación, implementación de aplicaciones web. Implica la adhesión a buenas prácticas, estándares de diseño, escalabilidad, rendimiento y seguridad con el objetivo de garantizar que un sitio o aplicación web cumpla con las expectativas de los usuarios, sea funcionalmente sólido, brinde una experiencia positiva y alcance los objetivos previstos. La calidad de desarrollo web se mide a través de métricas objetivas, como el tiempo de carga, número de bugs, número de vulnerabilidades y número de códigos duplicados.

Esta definición conceptual destaca que la calidad de desarrollo web comprende el rendimiento, seguridad y la escalabilidad. La búsqueda de la calidad es esencial para mantener la competitividad en el entorno digital actual.

3.3 Población, muestra y muestreo

3.3.1 Población

82 archivos de la web

Según Sekaran & Bougie (2016) la población es un conjunto completo de elementos en el que se hace una investigación, en este sentido los 10 archivos de php constituyen la población de esta investigación.

Los criterios de selección de la población en función a la líneas de código escrito en cada archivo porque comprometen la calidad del desarrollo de un aplicativo web. Significa que es suficiente que exista un archivo php que contenga una línea de código que comprometa alguna funcionalidad de la web para ser seleccionado como objeto de estudio.

3.3.2 Muestreo

En el marco de esta investigación, se ha adoptado un enfoque exhaustivo que implica el análisis integral de todos los 10 archivos recopilados de la web. Según Fraenkel & Wallen (2028) detalla que exhaustiva implica el estudio de todos los individuos, casos o eventos en una población particular. En lugar de realizar un muestreo, la investigación se basa en la totalidad de la población disponible. Este enfoque se ha seleccionado con el objetivo de abordar exhaustivamente de cómo influye el uso de ChatGPT en la calidad de desarrollo

de una web en la empresa Investigación S.A.C en 2023, utilizando todos los datos disponibles para obtener una comprensión completa sobre la calidad del desarrollo web.

La utilización del 100% de los archivos asegura una cobertura completa de la población seleccionada, permitiendo una evaluación integral. Al no realizar un muestreo y utilizar todos los archivos disponibles, se busca evitar posibles sesgos que podrían surgir de una muestra más pequeña y garantizar la representatividad total de la población. Además se minimiza la pérdida de información, dado que la investigación está orientada a obtener una comprensión detallada de los archivos y sus características porque comprometen a la calidad del desarrollo web.

3.3.3 Unidad de análisis

Según Hancock et al. (2021), la unidad de análisis es referente a la entidad o fenómeno que se examina en profundidad durante el estudio. Por tal motivo la unidad de análisis de esta investigación fueron los archivos que componen la web porque han sido examinados de manera minuciosa.

3.4 Técnicas e instrumentos de recolección de datos

La técnica de recopilación de datos es fichaje, a través de la visualización de los resultados que fueron registrados los datos de grupo control y grupo experimental en las fichas de registro. Las herramientas que se han usado previa validación fueron: prueba de calidad de código con SonarQube y Sonar Scanner (test). La herramienta da como resultado un reporte (Ver anexo 1).

Con la técnica y los instrumentos mencionados se pudo registrar el número de vulnerabilidades, Code Smells, Bugs, código duplicado y Security Hotspots que implican en la calidad de desarrollo web.

3.4.1 Validez del instrumento

La constatación de esta investigación fue a través del “juicio de los expertos” que es un instrumento que sirve para recopilar la evaluación de los expertos en esta investigación, Lo expertos indicaron que la herramienta es válido para realizar el estudio (Ver anexo 2)

Tabla 1: Resumen de validación de expertos

Indicador	Experto 1	Experto 2	Validez
Tiempo promedio de respuesta del servidor	4	4	Alto nivel
Puntuación del rendimiento de Page Speed	4	4	Alto nivel
Número de solicitudes de HTTP por página	4	4	Alto nivel
Número de vulnerabilidades identificados	4	4	Alto nivel
Números de Bugs identificados	4	4	Alto nivel
Número de Security Hotspots	4	4	Alto nivel
Número de Code Smell	4	4	Alto nivel
Número de código duplicado	4	4	Alto nivel

Fuente: Elaboración propia.

3.5. Procedimientos

El primer paso para la obtención de los datos fue instalar SonarQube en localhost. Los pasos fueron: verificar si la computadora que se va usar para análisis tiene instalado Java V17 porque SonarQube requiere Java para ser ejecutado. El segundo paso fue descargar la versión estable de SonarQube de la web oficial (<https://www.sonarqube.org/>). En disco C de la computadora se creó una carpeta llamada Testing para descomprimir los archivos.

Luego se configuró la base de datos en MySQL y un usuario dedicado para el funcionamiento del software en localhost..

Para configurar SonarQube se debe de ubicar el archivo en la carpeta que se ha creado C:\Testing\sonarqube-10.2.1.78527\conf\sonar.properties y configura la conexión a tu base de datos.

Para analizar el proyecto con SonarScanner se debe de verificar que esté instalado dicho software. Como se estuvo comprobado que existe la instalación de SonarScanner se continúa con la configuración del proyecto.

El siguiente paso fue ejecutar el script .bat de inicio desde la carpeta C:\Testing\sonar-scanner-5.0.1.3006-windows\bin, es decir ejecutar desde CMD sonar-scanner.bat.

El archivo .bat ejecutado dio como resultado que se inició el análisis del proyecto indicando que se puede acceder desde <http://localhost:9000>. Importante

que dentro del archivo de sonar-project.properties debe estar configurado de la siguiente manera: grupo control y grupo experimental respectivamente

Proyecto 1:

```
sonar.projectKey=ester-xicota_1.02
sonar.projectName=ester-xicota_1.02
sonar.projectVersion=1.0
sonar.sources=C:/Testing/ester-xicota_1.02
sonar.languages=php
```

Proyecto 2:

```
sonar.projectKey=desafios_data
sonar.projectName=desafios_data
sonar.projectVersion=1.0
sonar.sources=C:/Testing/desafios_data
sonar.languages=php
```

```
sonar.projectName=ester-xicota_1.02
sonar.projectVersion=1.0
sonar.sources=C:/Testing/ester-xicota_1.02
sonar.languages=php

sonar.projectKey=desafios_data
sonar.projectName=desafios_data
sonar.projectVersion=1.0
sonar.sources=C:/Testing/desafios_data
sonar.languages=php
```

Para continuar con el procedimiento de recopilación de datos se accede a la carpeta de control C:\Testing\ester-xicota_1.02 se abre el proyecto desde CMD y se ejecuta el comando sonar-scanner, luego se realiza el mismo proceso para el grupo experimental. La ejecución del comando comenzó a recopilar información sobre los proyectos y enviar a la instancia de SonarQube.

Para la generación de reportes en SonarQube se accede a la Interfaz de usuario desde la ruta <http://localhost:9000> y se inicia la sesión con credenciales que se crearon durante la instalación. Luego se debe seleccionar el grupo control

para y se empieza a registrar los datos en la ficha de registro. Se repite el proceso para el grupo experimental.

El procedimiento se llevó de la siguiente manera: En la ficha de observación se registró el reporte que arrojó el software SonarQube de cada una de las aplicaciones webs desarrolladas con el uso y sin el uso de chatGPT en el proceso de codificación.

3.6 Método de análisis de datos

En el proceso de análisis de la información de los datos recopilados, se empleó la estadística descriptiva usando SPSS. Para el análisis se usó Mann-Whitney porque los datos no son paramétricos y las pruebas son muestras independientes. Para realizar el análisis de datos primero se identificaron las variables y se crearon los grupos control y experimental. Para proceder el análisis se acude al menú de SPSS analyze, realizando una búsqueda nonparametric Tests, independent Samples, y se elige la prueba U Mann Whitney y finalmente click en Ok. y se ejecutó el análisis. Este enfoque metodológico permitió derivar conclusiones que se basan exclusivamente en la información recopilada mediante los instrumentos utilizados. Para llevar a cabo el análisis de la información y los datos reunidos, se seguirá un enfoque que toma en consideración el perfil de uso de los datos.

La siguiente etapa involucra la codificación de la información recopilada con el propósito de llevar a cabo un análisis de las variables independientes y dependientes. Cada indicador recopilado fue sometido a una evaluación estadística exhaustiva. En este contexto, se adoptará un enfoque cuantitativo para el análisis de los datos, el cual comprende tanto un análisis estadístico como no paramétrico. La presentación final de los datos se realizó mediante diseños y tableros que ofrecerán una representación clara y objetiva.

Durante la fase de análisis descriptivo, se utilizarán tablas y gráficos de frecuencia para proporcionar una representación visual efectiva de los datos recopilados. Este enfoque integral garantizará una interpretación profunda y rigurosa de la información, contribuyendo así a la validez y robustez de los resultados obtenidos.

3.7 Aspectos éticos

Este estudio se ejecutó en estricto cumplimiento con las pautas establecidas por la Resolución Rectoral RVI N°062-2023-VI-UCV, que delineó los lineamientos esenciales para el desarrollo de investigaciones, alineándose con las políticas y procedimientos fundamentados por la universidad. Asimismo, se tuvo en cuenta la Resolución de Consejo Universitario N°0470-2022/UCV, la cual aborda las normativas destinadas a promover la identidad científica y ética, asegurando la rigurosidad científica, la protección de los participantes y la salvaguarda de la propiedad intelectual.

Se dio especial atención a la normativa RVI N°066-2023-VI-UCV, cuyo propósito es establecer las directrices para la planificación y desarrollo de las actividades académicas de las Experiencias Curriculares (ECs) del Programa de Investigación Formativa de la Universidad César Vallejo. Su ámbito de aplicación cubre todas las ECs del programa en modalidades de pregrado, posgrado y UCV Virtual, abarcando la preparación, ejecución, evaluación y cierre de los semestres académicos de la UCV. Se respetaron y reconocieron las ideas de los autores mencionados en las referencias bibliográficas, aplicando las normas APA, séptima edición, para su debida citación.

En cuanto a los datos utilizados en este estudio, se adquirieron de manera confiable y confidencial de la empresa de desarrollo de software. Se garantiza que la información fue tratada en su totalidad conforme a los protocolos establecidos. El gerente de la empresa destacó que esta información sólo deberá ser empleada con fines académicos y no debe publicarse ni difundirse sin la previa autorización del autor, preservando así los intereses de la empresa.

IV. RESULTADOS

Los valores de los resultados con referente a la vulnerabilidad oscilan entre 0 y 1,203,500. Teniendo en cuenta que la mediana (50%) es 1,000, significa que los valores son bajos, lo que podría indicar que la cantidad de vulnerabilidades detectadas es nula.

En tanto los valores de los Bugs oscilan entre -144 y 1,193,500. Un valor negativo indica un cambio en la disminución de errores. La mediana es menor que el máximo, significa que existe diferencia de errores entre el grupo control y experimental, el número de bugs en el grupo experimental es bajo.

Los valores de hotspost oscilan entre -150 y 1,196,500. Similar que con los bugs, un valor negativo que indica un cambio en la reducción de problemas. La media y la mediana son mucho menores que el máximo, lo que evidencia que existen menos hotspots en el grupo experimental que en el grupo control.

Ahora si se ingresa al detalle de los valores de CodeSmell oscilan entre -352 y 2,163,500. Un valor negativo que demuestra una mejora en la calidad del código. Dado que la mediana es 725, y el valor máximo es significativamente alto, esto demuestra que el grupo control presenta un alto número de "code smells".

Finalmente los valores de CodeDuplicado oscilan entre -1460 y 1,111,000. Un valor negativo que representa la disminución de la duplicación de código en el grupo experimental. La mediana es 144, lo que evidencia que la cantidad de código duplicado es baja.

4.3 Prueba de hipótesis

En esta investigación, se comparó la calidad de desarrollo web que se hizo con el uso de chatGPT y la calidad de desarrollo web sin el uso del chatGPT.

Estadísticos de prueba

	Vulnerabilidad Bugs		Hotspost	CodeSmell	CodeDuplicado
U de Mann-Whitney	768.500	758.500	761.500	732.500	676.000
W de Wilcoxon	1.203.500	1.193.500	1.196.500	2.163.500	1.111.000
Z	0	-144	-150	-352	-1.460

Sig. asin. (bilateral)	1.000	0.885	0.880	0.725	0.144
a Variable de agrupación: Grupo					

Vulnerabilidad: los resultados de la prueba U de Mann-Whitney muestran que no hay evidencia de que la cantidad de vulnerabilidades sea diferente entre los dos grupos. La hipótesis nula no se puede rechazar, lo que indica que la cantidad de vulnerabilidades es lo mismo en ambos grupos.

Bugs: los resultados de la prueba U de Mann-Whitney muestran que hay evidencia de la cantidad de bugs entre los dos grupos. La hipótesis nula se rechaza, lo que indica que hay más bugs en el grupo control que en el grupo experimental.

Hotspot: los resultados de la prueba U de Mann-Whitney muestran que hay evidencia de que la cantidad de hotspots es diferente entre los dos grupos. La hipótesis nula se rechaza, lo que indica que hay más hotspots en el grupo control que en el grupo experimental.

Code Smell: los resultados de la prueba U de Mann-Whitney muestran que existe una diferencia de la cantidad de code smells entre los dos grupos. La hipótesis nula se rechaza, lo que indica es que la cantidad de code smells es mayor en el grupo control que en el grupo experimental.

Code duplicado: los resultados de la prueba U de Mann-Whitney muestran que hay evidencia de que la cantidad de código duplicado es diferente entre los dos grupos. La hipótesis nula se rechaza, lo que indica que existe más código duplicado en el grupo control que en el grupo experimental.

La hipótesis nula (H0) fue que el grupo experimental es mayor que el grupo control, por la cantidad de errores y proceso que genera durante el desarrollo. La hipótesis alternativa (H1) fue que el grupo experimental es menor que el grupo control, porque a medida que mejora los procesos de desarrollo bajan los números de errores en los archivos que componen la web.

El estudio incluyó a dos proyectos de desarrollo web que se compone de sus propios archivos, el grupo control arrojó de 53 archivos y el grupo experimental 29

archivos. Cada uno de estos archivos fueron analizados a detalle con SonarQube para evitar los sesgos.

H0 = El grupo de experimental es mayor que grupo control

H1 = El grupo de experimental es menor que el grupo control

Prueba de Mann-Whitney

	Grupo	N	Rango promedio	Suma de rangos
Vulnerabilidad	control	53	41.50	2199.50
	experimental	29	41.50	1203.50
	Total	82		
Bugs	control	53	41.69	2209.50
	experimental	29	41.16	1193.50
	Total	82		
Hotspot	control	53	41.63	2206.50
	experimental	29	41.26	1196.50
	Total	82		
CodeSmell	control	53	40.82	2163.50
	experimental	29	42.74	1239.50
	Total	82		
CodeDuplicado	control	53	43.25	2292.00
	experimental	29	38.31	1111.00
	Total	82		

V. DISCUSIÓN

En base a los resultados para la variable dependiente de **calidad de desarrollo web**, se obtuvo una mejora del 99.12% en base a la comparativa entre las medias del pretest y el posttest para este caso. Al comprobar las similitudes con la Investigación de Ozturk et al.: El resultado contradice los hallazgos quienes afirmaron que ChatGPT tenía un 91% de falsos positivos en comparación con los analizadores tradicionales. La falta de diferencia significativa podría sugerir que, en este estudio, ChatGPT no mostró una tendencia marcada hacia falsos positivos. Con referente a los bugs, el resultado de la Prueba de Mann-Whitney: Hay evidencia de que la cantidad de bugs es diferente entre los dos grupos ($p < 0.05$). Comparación con la Investigación de Ozturk et al.: Los resultados concuerdan con la advertencia de Ozturk et al. sobre la precisión actual de ChatGPT. La presencia significativa de bugs en el grupo control indica que ChatGPT tiene la corrección efectiva de errores. Y se verifican los casos de Hotspots, Code Smell y Code Duplicado: Resultados de la Prueba de Mann-Whitney: En hotspots y code smell, hay evidencia de que la cantidad es diferente entre los dos grupos ($p < 0.05$). En code duplicado, también hay evidencia de diferencia ($p < 0.05$). Comparación con la Investigación de Ozturk et al.: Los resultados respaldan la sugerencia de explorar la combinación de analizadores tradicionales con ChatGPT. La detección significativa de hotspots, code smell y código duplicado en el grupo control resalta que la aplicación con chatGPT tiene mejores resultados.

La conexión entre los resultados cuantitativos de Wilcoxon y las afirmaciones del estudio de Yilmaz y Yilmaz (2023) respalda la idea de que el uso de ChatGPT influye positivamente en el pensamiento lógico y, por ende, mejora la calidad del código en términos de vulnerabilidades, bugs y hotspots. Estos hallazgos abren oportunidades para seguir explorando el papel de ChatGPT en el proceso de desarrollo de aplicaciones y software y su impacto en las habilidades cognitivas de los programadores.

La relación entre los resultados de la Prueba de Wilcoxon y la investigación de Fajkovic y Rundberg sugiere que, si bien la calidad del código puede ser similar entre Copilot y ChatGPT, la facilidad de uso en la creación inicial de sitios web podría inclinar la balanza a favor de ChatGPT en ciertos casos.

Estos hallazgos subrayan la importancia de evaluar no solo la calidad del código sino también la eficiencia y la facilidad de implementación en el desarrollo web.

Dentro del aspecto de **Escalabilidad de la web**, se pudo determinar que, al momento de realizar una comparativa entre el promedio del pretest con el del posttest, se obtuvo una mejora del 434.93%. La conexión entre los resultados de la Prueba de Wilcoxon y la capacidad de ChatGPT como herramienta de apoyo sugiere un escenario positivo en el que la integración de esta tecnología puede ser beneficiosa para el desarrollo de software, ofreciendo generación y explicación de código efectivas.

La conclusión de Al-Khiami y Jaeger (2023) sobre la importancia de adoptar un enfoque por fases en el desarrollo con ChatGPT ofrece una perspectiva valiosa que se relaciona con los resultados de la Prueba de Mann-Whitney.

Los resultados obtenidos en la Prueba de Wilcoxon brindan un marco para discutir aspectos específicos de la utilidad y aplicabilidad de ChatGPT en el contexto de desarrollo de software. Estos resultados, junto con las observaciones del investigador York, E. (2023), destacan la importancia de considerar aspectos organizativos y evaluativos de ChatGPT, especialmente en el ámbito de la Experiencia del Usuario (UX).

Los resultados de la Prueba de Wilcoxon proporcionan un contexto valioso para discutir el impacto de ChatGPT en el desarrollo de software, especialmente en relación con la automatización de tareas y la mejora de la calidad del código. La investigación de White, J., Hays, S., Fu, Q., Spencer-Smith, J., & Schmidt, D. C. (2023) y Backström, O., & Kihlert, A. (2023) destaca la importancia de considerar el papel de ChatGPT en las fases de ejecución de ingeniería de software.

Los resultados de la Prueba de Wilcoxon proporcionan una plataforma para explorar la eficiencia y el impacto potencial de ChatGPT en el ámbito de las pruebas de software y la programación, a la luz de las conclusiones de investigaciones previas. La eficacia en la ejecución de casos de prueba, como sugiere Lesbekkyzy, A. (2023), subraya el potencial de ChatGPT para abordar tareas específicas en el dominio de la calidad del software. La capacidad para

automatizar y mejorar el proceso de prueba puede tener implicaciones significativas en la eficiencia global del desarrollo.

Los resultados obtenidos de la Prueba de Wilcoxon proporcionan un contexto valioso para reflexionar sobre la calidad del código generado por ChatGPT, según lo destacado por la investigación de Liu et al. (2023). La discusión sobre la calidad del código generado por ChatGPT sugiere que puede haber un equilibrio delicado entre la capacidad para producir código funcional y la necesidad de garantizar que dicho código cumpla con estándares aceptables de calidad. Es fundamental evaluar cómo se ponderan estos factores en el contexto del desarrollo de software.

Los resultados de la Prueba de Wilcoxon proporcionan un marco para discutir la efectividad en la generación de código, especialmente en comparación con un estudio adicional realizado por Yetiştirilen et al. (2023). En este estudio, se evaluaron GitHub Copilot, Amazon CodeWhisperer y ChatGPT en términos de su capacidad para generar soluciones de código. Los resultados del estudio indican que ChatGPT tuvo la tasa de éxito más alta entre los evaluados, logrando generar soluciones de código correctas para el 65,2% de los problemas en el conjunto de datos de HumanEval. Este hallazgo destaca la eficacia de ChatGPT en la generación de código funcional y preciso.

En base a la **seguridad de la web**, se pudo determinar que hubo una mejora del 80% al momento de hacer una comparativa entre el promedio del pretest con el de posttest. Los resultados de la Prueba de Wilcoxon proporcionan una plataforma para reflexionar sobre las implicaciones pedagógicas de las herramientas de codificación de IA, especialmente a la luz de las inquietudes expresadas por los investigadores Lau & Guo (2023). La investigación y las reflexiones de Lau y Guo abren un diálogo valioso sobre el futuro de la educación en programación en un mundo donde las herramientas de codificación de IA desempeñan un papel central. La discusión puede servir como punto de partida para abordar los desafíos y oportunidades pedagógicos en un entorno en constante evolución.

Estos resultados sirven como punto de partida para discutir la integración de aplicaciones de inteligencia artificial, particularmente ChatGPT, en el proceso de aprendizaje de programadores, tal como lo indican los autores Waseem et al.

(2023). La idea de que la incorporación de instrumentos de inteligencia artificial, como ChatGPT, puede aportar beneficios sustanciales al proceso de desarrollo de software. La reducción en la vulnerabilidad y errores podría indicar un impacto positivo.

Aunque los resultados sugieren mejoras en la vulnerabilidad, errores y hotspots, la discusión se amplía a la eficiencia global y los posibles riesgos asociados. ¿En qué medida la eficiencia justifica el riesgo de depender en gran medida de las herramientas de IA? La preocupación expresada por Tomic sobre la pérdida de habilidades y el toque humano en el desarrollo de software subraya la necesidad de considerar cuidadosamente cómo estas herramientas afectan la participación activa y creativa de los desarrolladores. La investigación destaca la posibilidad de que los modelos de IA perpetúen sesgos existentes. ¿Cómo se abordan los desafíos de la justicia algorítmica y la equidad en el desarrollo de software? ¿Qué medidas se pueden tomar para mitigar los sesgos inherentes en los datos de entrenamiento?

En lo que respecta al indicador de **Rendimiento de la web** al momento de realizar una comparativa entre los promedios del pretest y el posttest, se pudo determinar que se obtuvo una mejora del 1,37%, dentro de esta discusión sobre la Eficiencia de GPT-4 en el Análisis de Código Estático (Szabó & Bilicki, 2023): Los resultados de la Prueba de Wilcoxon sugieren una diferencia significativa entre los grupos de control y experimental en términos de vulnerabilidad, errores y hotspots. Es fundamental profundizar en estos hallazgos, contextualizando el estudio de Szabó y Bilicki (2023) sobre el enfoque de seguridad de aplicaciones web con GPT-4. El estudio revela que GPT-4 logró una precisión del 88,76% en el análisis de código estático. Esto resalta las capacidades de interpretación de modelos de lenguaje grandes como GPT-4 en la identificación de vulnerabilidades y errores en el código.

Los resultados obtenidos de la Prueba de Wilcoxon revelan un análisis comparativo entre el grupo de control y el grupo experimental, enfocándose en la vulnerabilidad, errores (bugs) y hotspots. Estos hallazgos proporcionan una base sólida para abordar la eficacia de las herramientas de generación de código, específicamente ChatGPT, en comparación con Copilot, según el estudio de Hansson y Ellréus (2023).

La investigación de Hörnemalm destaca la necesidad de una gestión cuidadosa de la confianza de los desarrolladores junior en ChatGPT, resaltando la importancia de un enfoque educativo, un sistema de soporte robusto y la consideración de los riesgos potenciales asociados con la adopción de tecnologías emergentes en el desarrollo de software.

Los resultados de la Prueba de Mann-Whitney presentan datos valiosos sobre la percepción de vulnerabilidades, bugs y hotspots en el desarrollo de sitios web, y la investigación de Amri (2023) añade una perspectiva relevante al destacar la transformación de la industria mediante la integración de la inteligencia artificial (IA) en el desarrollo web. Esta integración ha impulsado mejoras significativas en diversas áreas, y su impacto continuo sugiere un cambio continuo en el panorama del desarrollo de sitios web.

Los resultados y observaciones presentados por Andersson y Marshall Olsson subrayan la importancia de un enfoque equilibrado y consciente al integrar ChatGPT en el desarrollo de software, teniendo en cuenta tanto las ventajas como las preocupaciones asociadas. Estas consideraciones son esenciales para maximizar el valor de ChatGPT y garantizar su implementación ética y efectiva en el ámbito del desarrollo de software.

La investigación presenta resultados de la prueba de Wilcoxon en tres categorías: Vulnerabilidad, Bugs y Hotspots, en los grupos de control y experimental. Además, se contextualiza la discusión con la preocupación de seguridad planteada por los autores Pa Pa, Y. M., Tanizaki, S., Kou, T., Van Eeten, M., Yoshioka, K., & Matsumoto, T. (2023), quienes alertan sobre el potencial de generación de malware y herramientas de ataque por parte de modelos de lenguaje como ChatGPT, Auto-GPT y text-davinci-003.

La investigación indica diferencias significativas entre los grupos de control y experimental en términos de "Vulnerabilidad", "Bugs" y "Hotspots". Al contextualizar estos hallazgos con el estudio de Zhang et al. (2023) sobre la reparación automatizada de programas utilizando ChatGPT, se abre la puerta a una discusión integral sobre la efectividad de ChatGPT en la mejora de la calidad del software.

El estudio de Xie et al. destaca que las pruebas unitarias realizadas con ChatGPT superan a EvoSuite, AthenaTest y A3Test en términos de cobertura de

líneas, sucursales y métodos focales. Esto sugiere un rendimiento superior en la generación de test unitarios en comparación con herramientas tradicionales.

Al integrar estos hallazgos con la información proporcionada por Sakib, Khan, & Karim (2023) sobre las capacidades de ChatGPT en la generación de soluciones de código, se abre una discusión rica en matices sobre las fortalezas y limitaciones de este modelo en el contexto de la programación. Se destaca que ChatGPT tiene un destacado rendimiento en la generación de soluciones de código, lo que sugiere su capacidad para asistir en tareas de programación y proporcionar soluciones viables. Por tanto los resultados de la prueba de Mann-Whitney, que indican diferencias entre los grupos de control y experimental en términos de vulnerabilidad, bugs y hotspots, se relacionan con las capacidades y limitaciones específicas de ChatGPT en la generación de soluciones de código. La integración de estas perspectivas proporciona una comprensión más completa de cómo este modelo puede afectar los flujos de trabajo de codificación.

Los resultados de la prueba de Wilcoxon indican que la introducción de ChatGPT en el grupo experimental está asociada con diferencias en la vulnerabilidad, bugs y hotspots. Esto sugiere un impacto potencial en la calidad del software y la corrección de errores.

Hallazgo: Tanto Copilot como ChatGPT entregan código de calidad similar, siendo más fácil crear un sitio web desde cero con ChatGPT y tiene un alto grado de éxito en la síntesis de programas de computadora y la explicación del código en lenguaje natural.

Otro hallazgo importante es que ChatGPT mejora la calidad del código y es efectivo en la detección de infracciones específicas. La capacidad de ChatGPT para abordar y resolver infracciones específicas destaca su utilidad en la mejora del código generado por los programadores.

VI. CONCLUSIONES

Se llegó a la conclusión de que el uso del ChatGPT tuvo un impacto positivo en la calidad de desarrollo web en la empresa Investigación donde se obtuvo una mejora del 99.12% y se pudo validar el tema de la hipótesis con un valor Z de -2,805 y una significancia menor al 0.05 de error establecido.

Se concluye que el uso de ChatGPT en el desarrollo de una web en la empresa Investigación S.A.C 2023 tuvo un efecto positivo en la escalabilidad de la web en donde mostró una mejora del 434.93% logrando validar el tema de la hipótesis con un valor Z de -2,805 y una significancia menor al 0.05 de error establecido.

En síntesis, el uso de ChatGPT en el desarrollo de una web en la empresa Investigación S.A.C 2023 tuvo un efecto positivo en la seguridad de la web el cual se obtuvo una mejora del 80% teniendo en cuenta que en la prueba de hipótesis se obtuvo un valor Z de -1.838 con una significancia menor al 0.05 de error establecido.

Por último, se concluye que el uso de ChatGPT en el desarrollo de una web en la empresa Investigación S.A.C 2023 tuvo un efecto positivo en el rendimiento de la web con una mejora del 1.37% sumado a que en las pruebas de hipótesis se obtuvo un valor Z de -1.780 con una significancia menor al 0.05 de error establecido.

VII. RECOMENDACIONES

Se recomienda el uso de ChatGPT en el desarrollo web porque influye en la calidad de desarrollo de una web. Queda demostrado según números de errores que presenta el grupo control versus experimental. El grupo control debe adoptar algunas de las prácticas exitosas implementadas en el grupo experimental, especialmente en términos de identificación y corrección de bugs, así como la reducción de code smells y código duplicado.

Para mantener altos estándares de calidad y prevenir posibles problemas en el futuro durante la escalabilidad del aplicativo web se recomienda que use de manera consciente ChatGPT. Importante considerar análisis adicionales para comprender mejor las causas subyacentes de las diferencias encontradas, lo que puede proporcionar información adicional para mejorar las prácticas de desarrollo web para ambos grupos.

Además, se recomienda usar ChatGPT para mejorar la seguridad del desarrollo de un aplicativo web, porque influye en el proceso de detección de errores en menor tiempo obteniendo así, un nivel de seguridad óptimo dentro de la plataforma web desarrollada por el usuario.

Para que el impacto del uso de ChatGPT en el rendimiento del desarrollo de un aplicativo web sea efectiva el programador debe tener conocimientos sólidos para poder ejecutar y monitorear el proceso de codificación.

REFERENCIAS

- Abbas, Q., Younus, W., Malik, S., & Hassan, M. H. (2023). Incorporating ChatGPT in Software Project Management. https://www.ieceem.com/researchpaper/ChatGPT_The_next_Frontier_in_Software_Project_Management.pdf
- Al-Khiami, M. I., & Jaeger, M. (2023). Leveraging ChatGPT in Android App Development: A Case Study on Supporting Novice Developers in Creating Successful Apps. <https://doi.org/10.20944/preprints202307.0660.v1>
- Amri, A. (2023). Artificial Intelligence (AI) Website Developer: Aesthetic Convenience For Web Designers And Graphic Designers Without Coding. *Jurnal Inovatif: Inovasi Teknologi Informasi dan Informatika*, 6(2), 121-125. <https://doi.org/10.32832/inovatif.v6i2.311>
- Andersson, M., & Marshall Olsson, T. (2023). ChatGPT as a Supporting Tool for System Developers: Understanding User Adoption. *diva2:1779321*
- Backström, O., & Kihlert, A. (2023). Code Quality and Large Language Models in Computer Science Education: Enhancing student-written code through ChatGPT. <https://www.diva-portal.org/smash/get/diva2:1779791/FULLTEXT01.pdf>
- Bale, A. S., Vada, Y. R., Oshiojum, B. E., Lakkineni, U. K., Rao, C., Venkatesh, K., & Rani, I. (2023). ChatGPT in Software Development: Methods and Cross-Domain Applications. *International Journal of Intelligent Systems and Applications in Engineering*, 11(9s), 636-643.
- Biswas, S. (2023). Role of ChatGPT in Computer Programming.: ChatGPT in Computer Programming. *Mesopotamian Journal of Computer Science*, 2023, 8-16. <https://doi.org/10.58496/MJCSC/2023/002>
- Bourenkov, G. P., & Popov, A. N. (2006). A quantitative approach to data-collection strategies. *Acta Crystallographica Section D: Biological Crystallography*, 62(1), 58-64. <https://doi.org/10.1107/S0907444905033998>

- Capuñay Uceda, O. E., & Antón Perez, J. M. (2021). Influencia de SCRUM en los plazos de entrega y rendimiento en los proyectos de las asignaturas de Desarrollo de Software. *Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología*, (29), 36-42. <http://www.scielo.org.ar/pdf/ritet/n29/n29a05.pdf>
- Elbrächter, D., Perekrestenko, D., Grohs, P., & Bölskei, H. (2019). Deep neural network approximation theory. *arXiv preprint arXiv:1901.02220*.
- Fajkovic, E., & Rundberg, E. (2023). The Impact of AI-generated Code on Web Development: A Comparative Study of ChatGPT and GitHub Copilot. <https://www.diva-portal.org/smash/get/diva2:1769082/FULLTEXT01.pdf>
- Feng, Y., Vanam, S., Cherukupally, M., Zheng, W., Qiu, M., & Chen, H. (2023). Investigating Code Generation Performance of Chat-GPT with Crowdsourcing Social Data. In *Proceedings of the 47th IEEE Computer Software and Applications Conference* (pp. 1-10). <https://doi.org/10.24963/ijcai.2018/578>
- Fraenkel, J., Wallen, N., & Hyun, H. (2018). *How to design and evaluate research in education (10th ed.)*. McGraw-Hill. <https://eric.ed.gov/?id=ED323168>
- Ge, H., & Wu, Y. (2023). An Empirical Study of Adoption of ChatGPT for Bug Fixing among Professional Developers. *Innovation & Technology Advances*, 1(1), 21-29. <https://doi.org/10.61187/ita.v1i1.19>
- Hancock, D. R., Algozzine, B., & Lim, J. H. (2021). Doing case study research: A practical guide for beginning researchers. https://books.google.com.pe/books?hl=es&lr=&id=e7ILEAAQBAJ&oi=fnd&pg=PP1&dq=Doing+case+study+research:+A+practical+guide+for+beginning+researchers.&ots=5-q8ROST1h&sig=0qXgaruNUFG8OKC7dJ2wW9QyPs8&redir_esc=y#v=onepage&q=Doing%20case%20study%20research%3A%20A%20practical%20guide%20for%20beginning%20researchers.&f=false

- Hansson, E., & Ellréus, O. (2023). Code Correctness and Quality in the Era of AI Code Generation: Examining ChatGPT and GitHub Copilot. *diva2:1764568*
- Hörnemalm, A. (2023). ChatGPT as a Software Development Tool: The Future of Development. *diva2:1768339*
- Latapie, H., Kilic, O., Liu, G., Kompella, R., Lawrence, A., Sun, Y., ... & Thórisson, K. R. (2021). A metamodel and framework for artificial general intelligence from theory to practice. *Journal of Artificial Intelligence and Consciousness*, 8(02), 205-227. <https://doi.org/10.1142/S2705078521500119>
- Latapie, H., Kilic, O., Liu, G., Kompella, R., Lawrence, A., Sun, Y., ... & Thórisson, K. R. (2021). A metamodel and framework for artificial general intelligence from theory to practice. *Journal of Artificial Intelligence and Consciousness*, 8(02), 205-227. <https://doi.org/10.1142/S2705078521500119>
- Lau, S., & Guo, P. J. (2023). From "Ban It Till We Understand It" to "Resistance is Futile. *How University Programming Instructors Plan to Adapt as More Students Use AI Code Generation and Explanation Tools such as ChatGPT and GitHub Copilot*. <https://doi.org/10.1145/3568813.3600138>
- Liu, Y., Le-Cong, T., Widyasari, R., Tantithamthavorn, C., Li, L., Le, X. B. D., & Lo, D. (2023). Refining ChatGPT-Generated Code: Characterizing and Mitigating Code Quality Issues. *arXiv preprint arXiv:2307.12596*. <https://doi.org/10.48550/arXiv.2307.12596>
- Marcillo Sánchez, P. M., & Román Barrezueta, L. D. (2022). Análisis de la información generada para mantener la escalabilidad y persistencia del proceso de desarrollo de software. *Serie Científica de la Universidad de las Ciencias Informáticas*, 15(8), 193-227. <https://dialnet.unirioja.es/servlet/articulo?codigo=8955513>
- Mehrad, A., & Zangeneh, M. H. T. (2019). Comparison between qualitative and quantitative research approaches: Social sciences. *International Journal For*

Research In Educational Studies, Iran, 5(7), 1-7.
<https://orcid.org/0000-0003-4364-5709>

- Miller, C. J., Smith, S. N., & Pugatch, M. (2020). Experimental and quasi-experimental designs in implementation research. *Psychiatry research, 283*, 112452. <https://doi.org/10.1016/j.psychres.2019.06.027>
- Mishra, S. B., & Alok, S. (2022). Handbook of research methodology. <https://www.nkrgacw.org/nkr%20econtent/nutrition%20and%20dietetics/PG/II.M.Sc%20N&D/BookResearchMethodology.pdf>
- Neuman, S. B., & McCormick, S. (1995). *Single-subject experimental research: Applications for literacy*. Order Department, International Reading Association, 800 Barksdale Road, PO Box 8139, Newark, DE 19714-8139 (Book No. 128: \$11 members, \$16 nonmembers).
- Othman, A., Dhouib, A., & Nasser Al Jabor, A. (2023, July). Fostering websites accessibility: A case study on the use of the Large Language Models ChatGPT for automatic remediation. In *Proceedings of the 16th International Conference on Pervasive Technologies Related to Assistive Environments* (pp. 707-713). <https://dl.acm.org/doi/abs/10.1145/3594806.3596542>
- Ozturk, O. S., Ekmekcioglu, E., Cetin, O., Arief, B., & Hernandez-Castro, J. (2023, June). New tricks to old codes: can AI chatbots replace static code analysis tools?. In *Proceedings of the 2023 European Interdisciplinary Cybersecurity Conference* (pp. 13-18). <https://dl.acm.org/doi/abs/10.1145/3590777.3590780>
- Pa Pa, Y. M., Tanizaki, S., Kou, T., Van Eeten, M., Yoshioka, K., & Matsumoto, T. (2023, August). An attacker's dream? exploring the capabilities of chatgpt for developing malware. In *Proceedings of the 16th Cyber Security Experimentation and Test Workshop* (pp. 10-18). <https://doi.org/10.1145/3607505.3607513>

- Pargaonkar, S. Cultivating Software Excellence: The Intersection of Code Quality and Dynamic Analysis in Contemporary Software Development within the Field of Software Quality Engineering.
- Petracca, E. (2021). Embodying bounded rationality: from embodied bounded rationality to embodied rationality. *Frontiers in Psychology*, 12, 710607. <https://doi.org/10.3389/fpsyg.2021.710607>
- Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. OpenAI blog, 1(8), 9. <https://insightcivic.s3.us-east-1.amazonaws.com/language-models.pdf>
- Rahmaniar, W. (2023). Chatgpt for software development: Opportunities and challenges. DOI:10.36227/techrxiv.23993583.v1
- Sakib, F. A., Khan, S. H., & Karim, A. H. M. (2023). Extending the frontier of chatgpt: Code generation and debugging. *arXiv preprint arXiv:2307.08260*. <https://doi.org/10.48550/arXiv.2307.08260>
- Sánchez-Bautista, G., & Ramírez-Chávez, L. (2022). Amenazas de seguridad a considerar en el desarrollo de software. *XIKUA Boletín Científico de La Escuela Superior de Tlahuelilpan*, 10(19), 31-37. <https://repository.uaeh.edu.mx/revistas/index.php/xikua/article/view/8118>
- Schwarz, G., Christensen, T., & Zhu, X. (2022). Bounded Rationality, Satisficing, Artificial Intelligence, and Decision-Making in Public Organizations: The Contributions of Herbert Simon. *Public Administration Review*. <https://doi.org/10.1111/puar.13540>
- Sekaran, U., & Bougie, R. (2016). *Research methods for business: A skill building approach*. John Wiley & Sons. <https://books.google.es/books?hl=es&lr=&id=Ko6bCgAAQBAJ&oi=fnd&pg=PA19&dq=research+Methods+for+Business:+A+Skill-Building+Approach&ots=2DZQW2IWnS&sig=7X0DRhRUQU1-bJWJHmj64-ApZzo#v=onepa>

ge&q=research%20Methods%20for%20Business%3A%20A%20Skill-Building%20Approach&f=false

Szabó, Z., & Bilicki, V. (2023). A New Approach to Web Application Security: Utilizing GPT Language Models for Source Code Inspection. *Future Internet*, 15(10), 326. <https://doi.org/10.3390/fi15100326>

Talabek, A. ARTIFICIAL AI IN TEST AUTOMATION: SOFTWARE TESTING OPPORTUNITIES WITH OPENAI TECHNOLOGY-CHATGPT. <https://doi.org/10.47344/sdubnts.v62i1.912>

Tian, H., Lu, W., Li, T. O., Tang, X., Cheung, S. C., Klein, J., & Bissyandé, T. F. (2023). Is ChatGPT the Ultimate Programming Assistant--How far is it?. *arXiv preprint arXiv:2304.11938*. <https://arxiv.org/abs/2304.11938>

Tosic, D. (2023). Artificial Intelligence-driven web development and agile project management using OpenAI API and GPT technology: A detailed report on technical integration and implementation of GPT models in CMS with API and agile web development for quality user-centered AI chat service experience. id: diva2:1764392

Veloz-Remache, G. D., Menéndez-Verdecia, J. A., & Aguilar-Moncayo, L. N. (2021). Mejores prácticas de calidad en el desarrollo de software integradas al conocimiento de la ingeniería. *Polo del Conocimiento*, 6(1), 656-668. <https://polodelconocimiento.com/ojs/index.php/es/article/view/2170>

Wang, P., & Goertzel, B. (Eds.). (2012). *Theoretical foundations of artificial general intelligence* (Vol. 4). Springer Science & Business Media. https://books.google.com.pe/books?hl=es&lr=&id=9B00vTQlh5sC&oi=fnd&pg=PR3&dq=Theoretical+foundations+of+artificial+general+intelligence&ots=0KhIxBLmF&sig=BjdqgfCBrZkxKQ0FfBAo-DD_8Bk&redir_esc=y#v=onepage&q=Theoretical%20foundations%20of%20artificial%20general%20intelligence&f=false

Waseem, M., Das, T., Ahmad, A., Fehmideh, M., Liang, P., & Mikkonen, T. (2023). Using ChatGPT throughout the Software Development Life Cycle by Novice

- Developers. *arXiv preprint arXiv:2310.13648*.
<https://doi.org/10.48550/arXiv.2310.13648>
- White, J., Hays, S., Fu, Q., Spencer-Smith, J., & Schmidt, D. C. (2023). Chatgpt prompt patterns for improving code quality, refactoring, requirements elicitation, and software design. *arXiv preprint arXiv:2303.07839*.
<https://doi.org/10.48550/arXiv.2303.07839>
- Xie, Z., Chen, Y., Zhi, C., Deng, S., & Yin, J. (2023). ChatUniTest: a ChatGPT-based automated unit test generation tool. *arXiv preprint arXiv:2305.04764*. <https://doi.org/10.48550/arXiv.2305.04764>
- Yetiştirilen, B., Özsoy, I., Ayerdem, M., & Tüzün, E. (2023). Evaluating the Code Quality of AI-Assisted Code Generation Tools: An Empirical Study on GitHub Copilot, Amazon CodeWhisperer, and ChatGPT. *arXiv preprint arXiv:2304.10778*. <https://doi.org/10.48550/arXiv.2304.10778>
- Yilmaz, R., & Yilmaz, F. G. K. (2023). Augmented intelligence in programming learning: Examining student views on the use of ChatGPT for programming learning. *Computers in Human Behavior: Artificial Humans*, 1(2), 100005. <https://www.sciencedirect.com/science/article/pii/S2949882123000051>
- York, E. (2023, October). Evaluating ChatGPT: Generative AI in UX Design and Web Development Pedagogy. In *Proceedings of the 41st ACM International Conference on Design of Communication* (pp. 197-201). <https://doi.org/10.1145/3615335.3623035>
- Zhang, Q., Zhang, T., Zhai, J., Fang, C., Yu, B., Sun, W., & Chen, Z. (2023). A Critical Review of Large Language Model on Software Engineering: An Example from ChatGPT and Automated Program Repair. *arXiv preprint arXiv:2310.08879*. <https://doi.org/10.48550/arXiv.2310.08879>

ANEXOS

Anexo 1 Instrumento de recolección de datos

Ficha de observación del grupo de control y grupo experimental

Fecha de observación:	00-00-0000
Observador:	Darwin Robles
Lugar:	Oficina principal de la empresa
Hora de inicio:	00:00:00
Descripción de la observación:	
1.- Objetivo de la observación: El objetivo de esta dimensión es conocer el rendimiento de la web	
Detalla de la observación:	
Descripción:	Cantidad:
Tiempo promedio de respuesta del servidor	
Puntuación del rendimiento de Page Speed	
Número de solicitudes de HTTP por página	
2.- Objetivo de la observación: El objetivo de esta dimensión es medir el nivel de seguridad de la aplicación web	
Detalla de la observación:	
Descripción:	Cantidad:
Número de vulnerabilidades identificados	
Números de Bugs identificados	
Número de Security Hotspots	
3.- Objetivo de la observación: El objetivo de esta dimensión es medir la escalabilidad de la aplicación web con menor dificultad.	
Detalla de la observación:	
Descripción:	Cantidad:
Número de Code Smell	
Número de código duplicado	

Anexo 2. Evaluación por juicio de expertos

Respetado juez: Usted ha sido seleccionado para evaluar el instrumento “**Ficha de observación del grupo de control**”. La evaluación del instrumento es de gran relevancia para lograr que sea válido y que los resultados obtenidos a partir de éste sean utilizados eficientemente; aportando al quehacer del desarrollo web con ChatGPT. Agradecemos su valiosa colaboración.

1. Datos generales del juez

Nombre del juez:	Juan Brues Chumpe Agosto
Grado profesional:	Maestría () Doctor (x)
Área de formación académica:	Clínica () Social () Educativa (x) Organizacional ()
Áreas de experiencia profesional:	Educación / Ingeniería
Institución donde labora:	Universidad Privada César Vallejo
Tiempo de experiencia profesional en el área:	2 a 4 años () Más de 5 años (x)
Experiencia en Investigación (si corresponde)	>6 años

2. Propósito de la evaluación:

Validar el contenido del instrumento, por juicio de expertos.

3. Datos

Nombre de la Prueba:	Ficha de observación del grupo de control y experimental
Autor (a):	Robles Luján, Darwin Edwin
Procedencia:	Perú
Administración:	Autocumplimentado
Tiempo de aplicación:	10 minutos
Ámbito de aplicación:	Desarrollo web en la empresa Investigación S.A.C.
Significación:	Nivel de significancia: 0.05

4. Soporte teórico

Variable	Dimensiones	Definición
Calidad de desarrollo web	Rendimiento	La "Calidad de Desarrollo Web" se refiere a la excelencia, eficiencia y cumplimiento de estándares en la creación de una web. Incluye dimensiones como el rendimiento, seguridad, la escalabilidad y la ausencia de errores o
	Seguridad	
Categoría	Calificación	Indicador
CLARIDAD El ítem se comprende fácilmente, es decir, su sintáctica y semántica son adecuadas.	1. No cumple con el criterio	El ítem no es claro.
	2. Bajo Nivel	El ítem requiere bastantes modificaciones o una modificación muy grande en el uso de las palabras de acuerdo con su significado o por la ordenación de estas.
	3. Moderado nivel	Se requiere una modificación muy específica de algunos de los términos del ítem.
	4. Alto nivel	El ítem es claro, tiene semántica y sintaxis adecuada.
COHERENCIA El ítem tiene relación lógica con la dimensión o indicador que está midiendo.	1. Totalmente en desacuerdo (no cumple con el criterio)	El ítem no tiene relación lógica con la dimensión.
	2. Desacuerdo (bajo nivel de acuerdo)	El ítem tiene una relación tangencial /lejana con la dimensión.
	3. Acuerdo (moderado nivel)	El ítem tiene una relación moderada con la dimensión que se está midiendo.
	4. Totalmente de acuerdo (alto nivel)	El ítem se encuentra está relacionado con la dimensión que está midiendo.
RELEVANCIA El ítem es esencial o importante, es decir debe ser incluido.	1. No cumple con el criterio	El ítem puede ser eliminado sin que se vea afectada la medición de la dimensión.
	2. Bajo Nivel	El ítem tiene alguna relevancia, pero otro ítem puede estar incluyendo lo que mide éste.
	3. Moderado nivel	El ítem es relativamente importante.
	4. Alto nivel	El ítem es muy relevante y debe ser incluido.
		problemas técnicos.

5. Presentación de instrucciones para el juez:

A continuación, a usted presento la ficha de observación del grupo control y experimental elaborado por Darwin Robles Luján en el año 2023. De acuerdo con los siguientes indicadores califique cada uno de los ítems según corresponda:

Leer con detenimiento los ítems y calificar en una escala de 1 a 4 su valoración, así como solicitamos brinde sus observaciones que considere pertinente

1. *No cumple con el criterio*
2. *Bajo Nivel*
3. *Moderado nivel*
4. *Alto nivel*

CERTIFICADO DE VALIDEZ POR JUICIO DE EXPERTOS QUE MIDE LA VARIABLE DE CALIDAD DESARROLLO WEB

Cada pregunta consta de una afirmación relacionada con cada dimensión de la variable y deberá evaluarla en una escala de tipo Likert de cinco puntos, donde 1 significa "Totalmente en desacuerdo" y 5 significa "Totalmente de acuerdo".

Dimensiones del instrumento: ficha de observación de desarrollo web

- **Primera dimensión: rendimiento**
- **Objetivos de la Dimensión:** El objetivo de esta dimensión es conocer el rendimiento de la web.

Indicadores	Ítem	Claridad				Coherencia				Relevancia				Observaciones/ Recomendaciones	
		1	2	3	4	1	2	3	4	1	2	3	4		
Tiempo promedio de respuesta del servidor	1. Tiempo de respuesta inferior a 200-300 milisegundos (Buena)														
	2. Tiempo de respuesta entre 300-500 milisegundos (regular)														
	3. Tiempo de respuesta mayor a 500 milisegundos (malo)														
Puntuación del rendimiento de Page Speed	4. entre 90-100 excelente 5. entre 50-89 bueno 6. entre 30-49 aceptable 7. entre 0 -29 Pobre														
Número de solicitudes de HTTP por página	8. número de HTTP request y response														

- **Segunda dimensión: seguridad**

- **Objetivos de la Dimensión:** El objetivo de esta dimensión es medir el nivel de seguridad de la aplicación web.

Indicadores	Ítem	Claridad				Coherencia				Relevancia				Observaciones/ Recomendaciones
		1	2	3	4	1	2	3	4	1	2	3	4	
Número de vulnerabilidades identificados	9. número de críticos 10. número de graves 11. número de menores													
Números de Bugs identificados	12. número de Críticos 13. número de Graves 14. número de menores													
Número de Security Hotspots	15. número de Critical 16. número de High 17. número de Medium 18. número de Low 19. número de Info													

- **Tercera dimensión: escalabilidad**

- **Objetivos de la Dimensión:** El objetivo de esta dimensión es medir la escalabilidad de la aplicación web con menor dificultad.

Indicadores	Ítem	Claridad				Coherencia				Relevancia				Observaciones/ Recomendaciones
		1	2	3	4	1	2	3	4	1	2	3	4	
Número de Code Smell	20. número de Critical 21. número de Major 22. número de Minor 23. número de Info													
Número de código duplicado	24. número de Critical 25. número de Major 26. número de Minor 27. número de Info													


 Chumpé Agosto Juan Brues
 DNI: 44824114

Firma del evaluador
CIP: 158102

**REGISTRO NACIONAL DE
GRADOS ACADÉMICOS Y TÍTULOS PROFESIONALES**

Graduado	Grado o Título	Institución
CHUMPE AGESTO, JUAN BRUES LEE DNI 44824114	INGENIERO DE SISTEMAS Fecha de diploma: 03/12/2013 Modalidad de estudios: -	UNIVERSIDAD PRIVADA CÉSAR VALLEJO PERU
CHUMPE AGESTO, JUAN BRUES LEE DNI 44824114	BACHILLER EN INGENIERIA DE SISTEMAS Fecha de diploma: 17/09/2013 Modalidad de estudios: - Fecha matrícula: Sin información (***) Fecha egreso: Sin información (***)	UNIVERSIDAD PRIVADA CÉSAR VALLEJO PERU
CHUMPE AGESTO, JUAN BRUES LEE DNI 44824114	MAGISTER EN DOCENCIA UNIVERSITARIA Fecha de diploma: 29/02/16 Modalidad de estudios: PRESENCIAL Fecha matrícula: 15/03/2013 Fecha egreso: 17/07/2013	UNIVERSIDAD PRIVADA CÉSAR VALLEJO PERU
CHUMPE AGESTO, JUAN BRUES LEE DNI 44824114	DOCTOR EN GESTIÓN PÚBLICA Y GOBERNABILIDAD Fecha de diploma: 12/11/18 Modalidad de estudios: PRESENCIAL Fecha matrícula: 17/08/2015 Fecha egreso: 21/08/2018	UNIVERSIDAD PRIVADA CÉSAR VALLEJO PERU
CHUMPE AGESTO, JUAN BRUES LEE DNI 44824114	MAESTRO EN INGENIERÍA DE SISTEMAS CON MENCIÓN EN TECNOLOGÍAS DE LA INFORMACIÓN Fecha de diploma: 10/12/18 Modalidad de estudios: PRESENCIAL Fecha matrícula: 05/04/2014 Fecha egreso: 25/03/2018	UNIVERSIDAD PRIVADA CÉSAR VALLEJO PERU
CHUMPE AGESTO, JUAN BRUES LEE DNI 44824114	BACHILLER EN DERECHO Fecha de diploma: 25/06/21 Modalidad de estudios: PRESENCIAL Fecha matrícula: 11/04/2015 Fecha egreso: 15/08/2020	UNIVERSIDAD PRIVADA SAN JUAN BAUTISTA S.A.C. PERU
CHUMPE AGESTO, JUAN BRUES LEE DNI 44824114	ABOGADO Fecha de diploma: 29/09/21 Modalidad de estudios: PRESENCIAL	UNIVERSIDAD PRIVADA SAN JUAN BAUTISTA S.A.C. PERU

Ficha de observación del grupo experimental y control

Fecha de observación:	00-00-0000
Observador:	Darwin Robles
Lugar:	Oficina principal de la empresa
Hora de inicio:	00:00:00
Descripción de la observación:	
1.- Objetivo de la observación:	Rendimiento
Detalla de la observación: El objetivo de esta dimensión es conocer el rendimiento de la web.	
Descripción:	Cantidad:
Tiempo promedio de respuesta del servidor	
Puntuación del rendimiento de Page Speed	
Número de solicitudes de HTTP por página	
2.- Objetivo de la observación:	Seguridad
Detalla de la observación: El objetivo de esta dimensión es medir el nivel de seguridad de la aplicación web.	
Descripción:	Cantidad:

Número de vulnerabilidades identificados	
Números de Bugs identificados	
Número de Security Hotspots	
2.- Objetivo de la observación:	Escalabilidad
Detalla de la observación: El objetivo de esta dimensión es medir la escalabilidad de la aplicación web con menor dificultad.	
Descripción:	Cantidad:
Número de Code Smell	
Número de código duplicado	

Evaluación por juicio de expertos (2)

Respetado juez: Usted ha sido seleccionado para evaluar el instrumento “**Ficha de observación del grupo de control**”. La evaluación del instrumento es de gran relevancia para lograr que sea válido y que los resultados obtenidos a partir de éste sean utilizados eficientemente; aportando al quehacer del desarrollo web con ChatGPT. Agradecemos su valiosa colaboración.

1. Datos generales del juez

Nombre del juez:	Abraham Rafael Sáenz Apari
Grado profesional:	Maestría (x) Doctor ()
Área de formación académica:	Clínica () Social () Educativa (x) Organizacional ()
Áreas de experiencia profesional:	Educación / Ingeniería
Institución donde labora:	Universidad Privada César Vallejo
Tiempo de experiencia profesional en el área:	2 a 5 años (x) Más de 5 años ()
Experiencia en Investigación (si corresponde)	

2. Propósito de la evaluación:

Validar el contenido del instrumento, por juicio de expertos.

Nombre de la Prueba:	Ficha de observación del grupo de control y experimental
Autor (a):	Robles Luján, Darwin Edwin
Procedencia:	Perú
Administración:	Autocumplimentado
Tiempo de aplicación:	10 minutos
Ámbito de aplicación:	Desarrollo web en la empresa Investigación S.A.C.
Significación:	Nivel de significancia: 0.05

3. Datos

Soporte teórico

Variable	Dimensiones	Definición
----------	-------------	------------

Categoría	Calificación	Indicador
CLARIDAD El ítem se comprende fácilmente, es decir, su sintáctica y semántica son adecuadas.	1. No cumple con el criterio	El ítem no es claro.
	2. Bajo Nivel	El ítem requiere bastantes modificaciones o una modificación muy grande en el uso de las palabras de acuerdo con su significado o por la ordenación de estas.
	3. Moderado nivel	Se requiere una modificación muy específica de algunos de los términos del ítem.
	4. Alto nivel	El ítem es claro, tiene semántica y sintaxis adecuada.
COHERENCIA El ítem tiene relación lógica con la dimensión o indicador que está midiendo.	1. Totalmente en desacuerdo (no cumple con el criterio)	El ítem no tiene relación lógica con la dimensión.
	2. Desacuerdo (bajo nivel de acuerdo)	El ítem tiene una relación tangencial /lejana con la dimensión.
	3. Acuerdo (moderado nivel)	El ítem tiene una relación moderada con la dimensión que se está midiendo.
	4. Totalmente de acuerdo (alto nivel)	El ítem se encuentra está relacionado con la dimensión que está midiendo.
RELEVANCIA El ítem es esencial o importante, es decir debe ser incluido.	1. No cumple con el criterio	El ítem puede ser eliminado sin que se vea afectada la medición de la dimensión.
	2. Bajo Nivel	El ítem tiene alguna relevancia, pero otro ítem puede estar incluyendo lo que mide éste.
	3. Moderado nivel	El ítem es relativamente importante.
	4. Alto nivel	El ítem es muy relevante y debe ser incluido.
Calidad de desarrollo web	Rendimiento	La "Calidad de Desarrollo Web" se refiere a la excelencia, eficiencia y cumplimiento de estándares en la creación de una página web. Incluye aspectos como el rendimiento, seguridad, la escalabilidad y la ausencia de errores o problemas técnicos.
	Seguridad	
	Escalabilidad	

4. Presentación de instrucciones para el juez:

A continuación, a usted presento la ficha de observación del grupo control y experimental elaborado por Darwin Robles Luján en el año 2023. De acuerdo con los siguientes indicadores califique cada uno de los ítems según corresponda:

Leer con detenimiento los ítems y calificar en una escala de 1 a 4 su valoración, así como solicitamos brinde sus observaciones que considere pertinente

1. *No cumple con el criterio*
2. *Bajo Nivel*
3. *Moderado nivel*
4. *Alto nivel*

CERTIFICADO DE VALIDEZ POR JUICIO DE EXPERTOS QUE MIDE LA VARIABLE CALIDAD DESARROLLO WEB

Cada pregunta consta de una afirmación relacionada con cada dimensión de la variable y deberá evaluarla en una escala de tipo Likert de cinco puntos, donde 1 significa "Totalmente en desacuerdo" y 5 significa "Totalmente de acuerdo".

Dimensiones del instrumento: ficha de observación de desarrollo web

- **Primera dimensión: rendimiento**

- Objetivos de la Dimensión: El objetivo de esta dimensión es conocer el rendimiento de la web.

Indicadores	Ítem	Claridad				Coherencia				Relevancia				Observaciones/ Recomendaciones
		1	2	3	4	1	2	3	4	1	2	3	4	
Tiempo promedio de respuesta del servidor	1. Tiempo de respuesta inferior a 200-300 milisegundos (Buena)													
	2. Tiempo de respuesta entre 300-500 milisegundos (regular)													
	3. Tiempo de respuesta mayor a 500 milisegundos (malo)													
Puntuación del rendimiento de Page Speed	4. entre 90-100 excelente													
	5. entre 50-89 bueno													
	6. entre 30-49 aceptable													
	7. entre 0 -29 Pobre													

- **Segunda dimensión: seguridad**

- Objetivos de la Dimensión: El objetivo de esta dimensión es medir el nivel de seguridad de la aplicación web.

Indicadores	Ítem	Claridad				Coherencia				Relevancia				Observaciones/ Recomendaciones
		1	2	3	4	1	2	3	4	1	2	3	4	
Número de vulnerabilidades identificados	8. número de críticos 9. número de graves 10. número de menores													
Números de Bugs identificados	11. número de Críticos 12. número de Graves 13. número de menores													
Número de Security Hotspots	14. número de Critical 15. número de High 16. número de Medium 17. número de Low 18. número de Info													

▪ **Tercera dimensión: escalabilidad**

- **Objetivos de la Dimensión:** El objetivo de esta dimensión es medir la escalabilidad de la aplicación web con menor dificultad.

Indicadores	Ítem	Claridad				Coherencia				Relevancia				Observaciones/ Recomendaciones
		1	2	3	4	1	2	3	4	1	2	3	4	
Número de Code Smell	19. número de Critical 20. número de Major 21. número de Minor 22. número de Info													
Número de código duplicado	23. número de Critical 24. número de Major 25. número de Minor													

**REGISTRO NACIONAL DE
GRADOS ACADÉMICOS Y TÍTULOS PROFESIONALES**

Graduado	Grado o Título	Institución
SAENZ APARI, ABRAHAM RAFAEL DNI 10454966	BACHILLER EN INGENIERIA DE COMPUTACION Y SISTEMAS Fecha de diploma: 10/06/2005 Modalidad de estudios: - Fecha matrícula: Sin información (***) Fecha egreso: Sin información (***)	UNIVERSIDAD DE SAN MARTÍN DE PORRES PERU
SAENZ APARI, ABRAHAM RAFAEL DNI 10454966	INGENIERO DE COMPUTACION Y SISTEMAS Fecha de diploma: 17/02/2006 Modalidad de estudios: -	UNIVERSIDAD DE SAN MARTÍN DE PORRES PERU
SAENZ APARI, ABRAHAM RAFAEL DNI 10454966	MAGISTER EN ADMINISTRACION ESTRATEGICA DE EMPRESAS Fecha de diploma: 21/08/2013 Modalidad de estudios: - Fecha matrícula: Sin información (***) Fecha egreso: Sin información (***)	PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ PERU

Ficha de observación del grupo experimental y control

Fecha de observación:	00-00-0000
Observador:	Darwin Robles
Lugar:	Oficina principal de la empresa
Hora de inicio:	00:00:00
Descripción de la observación:	
1.- Objetivo de la observación:	Rendimiento
Detalla de la observación: El objetivo de esta dimensión es conocer el rendimiento de la web.	
Descripción:	Cantidad:
Tiempo promedio de respuesta del servidor	
Puntuación del rendimiento de Page Speed	
Número de solicitudes de HTTP por página	
2.- Objetivo de la observación:	Seguridad
Detalla de la observación: El objetivo de esta dimensión es medir el nivel de seguridad de la aplicación web.	
Descripción:	Cantidad:
Número de vulnerabilidades identificados	

Números de Bugs identificados	
Número de Security Hotspots	
2.- Objetivo de la observación:	Escalabilidad
Detalla de la observación: El objetivo de esta dimensión es medir la escalabilidad de la aplicación web con menor dificultad.	
Descripción:	Cantidad:
Número de Code Smell	
Número de código duplicado	

Anexo 3. Reporte de webs desarrollada sin chatGPT y con el uso de chatGPT respectivamente

☆ [ester-xicota_1.02](#) PUBLIC ✓ Passed

Last analysis: 17 days ago • 5.2k Lines of Code • PHP, CSS

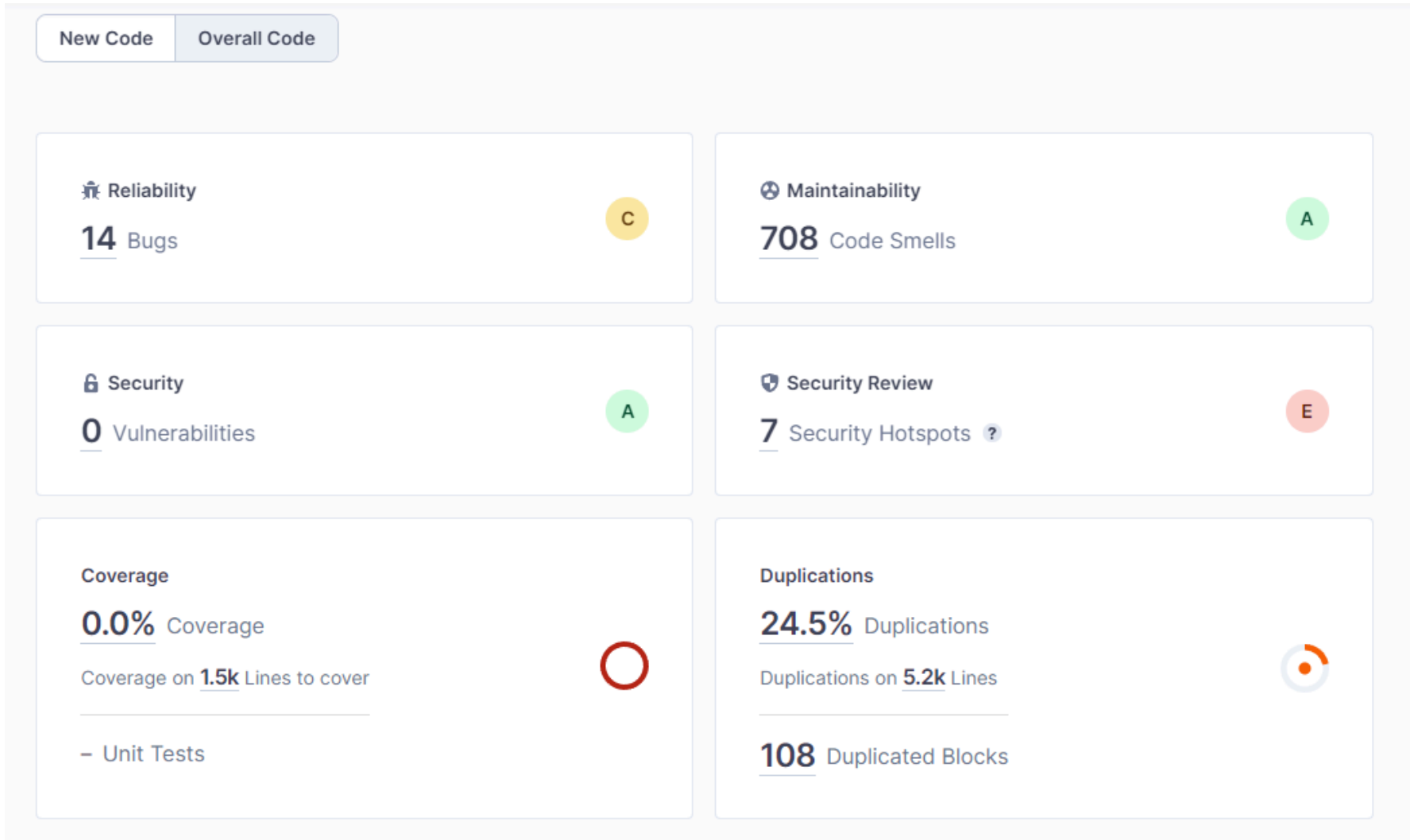
C 14	A 0	E 0.0%	A 708	○ 0.0%	○ 24.5%
Bugs	Vulnerabilities	Hotspots Reviewed	Code Smells	Coverage	Duplications

☆ [desafios_data](#) PUBLIC ✓ Passed

Last analysis: 1 day ago • 3.2k Lines of Code • CSS, PHP, ...

D 9	A 0	E 0.0%	A 184	○ 0.0%	○ 3.0%
Bugs	Vulnerabilities	Hotspots Reviewed	Code Smells	Coverage	Duplications

Reporte de SonarQube de web desarrollada sin intervención de chatGPT



Reporte de SonarQube de web desarrollada con la intervención de chatGPT

New Code

Overall Code

🔧 Reliability

9 Bugs

D

⚙️ Maintainability

184 Code Smells

A

🔒 Security

0 Vulnerabilities

A

🛡️ Security Review

2 Security Hotspots ?

E

Coverage

0.0% Coverage

Coverage on 627 Lines to cover

- Unit Tests

Duplications

3.0% Duplications

Duplications on 3.2k Lines

6 Duplicated Blocks

OPERACIONALIZACIÓN DE VARIABLE

VARIABLES	DEFINICIÓN CONCEPTUAL	DEFINICIÓN OPERACIONAL	DIMENSIONES	INDICADORES	ESCALA DE MEDICIÓN	TÉCNICAS / INSTRUMENTOS
VARIABLE INDEPENDIENTE Uso de chatGPT	El "Uso de la herramienta ChatGPT" se refiere a la incorporación y aplicación de un modelo de procesamiento de lenguaje natural (como ChatGPT) en el proceso de desarrollo de una web. Implica la utilización de ChatGPT para generar, mejorar o automatizar aspectos del contenido, diseño o funcionalidad del sitio web.	El "Uso de la herramienta ChatGPT" se medirá mediante el registro del número de interacciones de ChatGPT en el proceso de desarrollo web, con un prompt para generar el código durante el desarrollo web. La medición específica es desarrollo de web usando el chat y sin usar el chat.			Para medir el "Uso de la herramienta ChatGPT", utilizaremos una escala de medición que varía desde "No Uso" (0) hasta "Uso Extensivo" (1), donde 0 representa la ausencia de cualquier uso de ChatGPT en el desarrollo web y 1 indica un uso extensivo en múltiples aspectos del proceso de desarrollo.	
VARIABLE DEPENDIENTE Calidad de desarrollo web	La "Calidad de Desarrollo Web" se refiere a la excelencia, eficiencia y cumplimiento de estándares en la creación de una web web. Incluye dimensiones como el rendimiento, seguridad, la escalabilidad y la ausencia de errores o problemas técnicos.	La "Calidad de Desarrollo Web" se medirá mediante una combinación de indicadores, de vulnerabilidad, bugs, code smell, security hotspots, código duplicado, respuesta del servidor y solicitudes HTTP	Rendimiento	<ul style="list-style-type: none"> ● Tiempo promedio de respuesta del servidor ● Número de solicitudes de HTTP por página ● Puntuación de rendimiento según PageSpeed 	Para medir la "Calidad de Desarrollo Web", utilizamos escalas de intervalo (razón) para cada indicador, como número de errores de lógica por cada cada archivo, y una escala binaria de Sí/No para la aplicación de principios SOLID y patrones de diseño.	Ficha de observación del grupo experimental y grupo control
			Seguridad	<ul style="list-style-type: none"> ● Número de vulnerabilidades identificados 		

VARIABLES	DEFINICIÓN CONCEPTUAL	DEFINICIÓN OPERACIONAL	DIMENSIONES	INDICADORES	ESCALA DE MEDICIÓN	TÉCNICAS / INSTRUMENTOS
				<ul style="list-style-type: none"> ● Números de Bugs identificados ● Número de Security Hotspots 		
			Escalabilidad	<ul style="list-style-type: none"> ● Número de Code Smells (mala codificación) ● Número de código duplicado ● Capacidad de carga máxima del sitio sin degradación del rendimiento 		

Anexo 4. ASPECTOS ADMINISTRATIVOS

Recursos y Presupuesto

Tabla 1

Recursos y presupuesto

Clasificador de gasto MEF	Rubro	Cantidad	Costo Unitario	Costo Total
Bienes				
2.3.1.5.1.2	USB	2 unidades	S/ 20.00	S/ 40.00
2.3.1.5.1.3	Laptop Lenovo	1 unidad	S/ 4500.00	S/ 4500.00
Servicios				
2.3.1.5.1.2	Anillados	3 juegos	S/ 10.00	S/ 30.00
2.3.1.9.1.1	Impresión	150 hojas	S/ 0.30	S/ 45.00
2.3.2.7.2.2	Asesoría estadística	1 docente	S/ 1500.00	S/ 1500.00
2.3.2.1	Pasajes	20 pasajes	S/ 40.00	S/ 800.00
2.3.2.2.2.3	Internet	Por 6 meses	S/ 100.00	S/ 600.00
Total				S/ 7515.00

Elaboración propia







Financiamiento

Autofinanciado al 100%

Cronograma de ejecución

Actividades	Septiembre				Octubre				Noviembre				Diciembre				Enero	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Lineamientos para la ejecución del proyecto de investigación	x	x																
Búsqueda y acopio de información preliminar.		x	x	x														
Búsqueda bibliográfica inicial		x	x	x	x													
Presentación preliminar del marco teórico				x	x													
Presentación preliminar del proyecto					x	x												
Presentación de proyecto completo							x	x										
Diseño y validación de los instrumentos de investigación							x	x										
Reajuste y aprobación de plan de tesis.								x	x									
Trabajo de campo: contacto con la muestra - informantes								x	x									
Recolección de información								x	x									
Organización de la información									x	x								
Análisis de datos.									x	x								
Interpretación de datos										x	x							
Elaboración de conclusiones											x							
Redacción del informe											x	x						
Revisión y aprobación del informe													x	x	x	x		
Sustentación																		x

Anexo 5. Base de datos tabulado en SPSS y en Excel:

	 Grupo	 Vulnerabilidad	 Bugs	 Hotspot	 CodeSmell	 CodeDuplicado	var
1	1	0	0	0	47	42	
2	1	0	0	1	22	0	
3	1	0	1	0	2	0	
4	1	0	0	0	1	0	
5	1	0	0	0	1	0	
6	1	0	1	0	3	0	
7	1	0	1	0	3	0	
8	1	0	0	0	1	0	
9	1	0	1	0	1	0	
10	1	0	0	0	2	0	
11	1	0	0	0	3	0	
12	1	0	0	0	0	0	
13	1	0	0	1	5	0	
14	1	0	5	0	45	0	
15	1	0	1	0	7	0	
16	1	0	0	0	1	0	
17	1	0	0	0	3	0	
18	1	0	0	0	52	5	
19	1	0	0	0	1	0	
20	1	0	1	0	2	0	
21	1	0	0	0	51	5	
22	1	0	0	0	63	3	
23	1	0	0	0	3	0	
24	1	0	0	0	1	0	
25	1	0	1	0	1	0	
26	1	0	0	0	0	0	
27	1	0	0	0	0	0	
28	1	0	0	0	217	15	

Resultados sin GPT:

Archivo	name	# Vulnerabilidades	# Bugs	# Hotspots (security review)	# Code Smell	# Code Duplicado
1	customizer.php	0	0	0	47	42
2	front-page.php	0	0	1	22	0
3	404.php	0	1	0	2	0
4	archive.php	0	0	0	1	0
5	comments.php	0	0	0	1	0
6	content-category.php	0	1	0	3	0
7	content-none.php	0	1	0	3	0
8	content-page.php	0	0	0	1	0
9	content-search.php	0	1	0	1	0
10	content-single.php	0	0	0	2	0
11	custom-header.php	0	0	0	3	0
12	customizer.js	0	0	0	0	0
13	footer.php	0	0	1	5	0
14	functions.php	0	5	0	45	0
15	header.php	0	1	0	7	0
16	index.php	0	0	0	1	0
17	jetpack.php	0	0	0	3	0
18	page-moda-circular.php	0	0	0	52	5
19	navigation.js	0	0	0	1	0
20	page.php	0	1	0	2	0
21	page-formacion.php	0	0	0	51	5
22	single-servicios.php	0	0	0	63	3
23	search.php	0	0	0	3	0
24	sidebar.php	0	0	0	1	0
25	single.php	0	1	0	1	0
26	style-rtl.css	0	0	0	0	0
27	style.css	0	0	0	0	0
28	template-functions.php	0	0	0	217	45
29	template-tags.php	0	0	0	14	0
30	acordion-tabs.php	0	0	0	8	1
31	single-servicio.php	0	0	0	66	4
32	content-formacion.php	0	0	0	4	1
33	json-mailrealy.php	0	0	0	4	1
34	email-popup.php	0	0	2	8	1
35	author.php	0	0	0	3	0
36	boton-btn.php	0	0	0	3	0

37	boton-video.php	0	0	0	2	0
38	cat-cuatro.php	0	0	0	4	0
39	cat-dos.php	0	0	0	4	0
40	cat-tres.php	0	0	0	4	0
41	cat-uno.php	0	0	0	4	0
42	content-autores.php	0	1	0	4	0
43	content-autor.php	0	0	0	1	0
44	content-precio.php	0	0	0	4	0
45	content-registro.php	0	0	0	7	0
46	content-relacionados.php	0	0	0	7	0
47	content-video.php	0	1	0	4	0
48	home.php	0	0	0	4	0
49	input-paises.php	0	0	0	1	0
50	page-contacto.php	0	0	0	9	0
51	redes-sociales.php	0	0	3	2	0
52	sidebar-content.css	0	0	0	0	0
53	imagensubida.js	0	0	0	0	0

Resultados con GPT

Archivo	name	# Vulnerabilidades	# Bugs	# Hotspots (security review)	# Code Smell	# Code Duplicado
1	customizer.php	0	0	0	10	4
2	front-page.php	0	0	1	12	2
3	404.php	0	1	0	2	0
4	archive.php	0	0	0	2	0
5	comments.php	0	0	0	4	0
6	content-category.php	0	0	0	1	0
7	content-none.php	0	0	0	1	0
8	content-page.php	0	0	0	1	0
9	content-search.php	0	0	0	1	0
10	content.php	0	0	0	9	0
11	custom-header.php	0	0	0	3	0
12	customizer.js	0	0	0	1	0
13	footer.php	0	0	0	5	0
14	functions.php	0	5	0	16	0
15	header.php	0	1	1	14	0
16	index.php	0	0	0	1	0

17	jetpack.php	0	0	0	3	0
18	login-page.php	0	0	0	7	0
19	navigation.js	0	0	0	7	0
20	page.php	0	0	0	1	0
21	perfil-page.php	0	0	0	3	0
22	register-page.php	0	0	0	9	0
23	search.php	0	0	0	2	0
24	sidebar.php	0	0	0	1	0
25	single.php	0	0	0	1	0
26	style-rtl.css	0	0	0	9	0
27	style.css	0	1	0	15	0
28	template-functions.php	0	1	0	33	0
29	template-tags.php	0	0	0	12	0

Configuración para levantar SonarScanner:

```

8 #sonar.sourceEncoding=UTF-8
9 #proyecto Ester
10 sonar.projectKey=ester-xicota_1.02
11 sonar.projectName=ester-xicota_1.02
12 sonar.projectVesion=1.0
13 sonar.sources=C:/Testing/ester-xicota_1.02
14 sonar.languages=php
15
16 sonar.projectKey=desafios_data
17 sonar.projectName=desafios_data
18 sonar.projectVesion=1.0
19 sonar.sources=C:/Testing/desafios_data
20 sonar.languages=php
21
22 sonar.projectKey=storefront
23 sonar.projectName=storefront
24 sonar.projectVesion=1.0
25 sonar.sources=C:/Testing/storefront
26 sonar.languages=php
27
28

```

Comando para ejecutar SonarQube:

```
programa o archivo por lotes ejecutable.
C:\Testing\sonar-scanner-5.0.1.3006-windows>sonar-scanner.bat
INFO: Scanner configuration file: C:\Testing\sonar-scanner-5.0.1.3006-windows\bin\..\conf\sonar-scanner.properties
INFO: Project root configuration file: NONE
INFO: SonarScanner 5.0.1.3006
INFO: Java 17.0.7 Eclipse Adoptium (64-bit)
INFO: Windows 11 10.0 amd64
INFO: User cache: C:\Users\DARWIN\sonar\cache
ERROR: SonarQube server [http://localhost:9000] can not be reached
INFO: -----
INFO: EXECUTION FAILURE
INFO: -----
INFO: Total time: 0.283s
INFO: Final Memory: 3M/34M
INFO: -----
ERROR: Error during SonarScanner execution
org.sonarsource.scanner.api.internal.ScannerException: Unable to execute SonarScanner analysis
    at org.sonarsource.scanner.api.internal.IsolatedLauncherFactory.lambda$createLauncher$0(IsolatedLauncherFactory.java:85)
    at java.base/java.security.AccessController.doPrivileged(Unknown Source)
    at org.sonarsource.scanner.api.internal.IsolatedLauncherFactory.createLauncher(IsolatedLauncherFactory.java:74)
    at org.sonarsource.scanner.api.internal.IsolatedLauncherFactory.createLauncher(IsolatedLauncherFactory.java:70)
    at org.sonarsource.scanner.api.EmbeddedScanner.doStart(EmbeddedScanner.java:185)
    at org.sonarsource.scanner.api.EmbeddedScanner.start(EmbeddedScanner.java:123)
    at org.sonarsource.scanner.cli.Main.execute(Main.java:74)
    at org.sonarsource.scanner.cli.Main.main(Main.java:62)
Caused by: java.lang.IllegalStateException: Fail to get bootstrap index from server
    at org.sonarsource.scanner.api.internal.BootstrapIndexDownloader.getIndex(BootstrapIndexDownloader.java:42)
    at org.sonarsource.scanner.api.internal.JarDownloader.getScannerEngineFiles(JarDownloader.java:58)
    at org.sonarsource.scanner.api.internal.JarDownloader.download(JarDownloader.java:53)
    at org.sonarsource.scanner.api.internal.IsolatedLauncherFactory.lambda$createLauncher$0(IsolatedLauncherFactory.java:76)
    ... 7 more
Caused by: java.net.ConnectException: Failed to connect to localhost/[0:0:0:0:0:0:0:1]:9000
    at org.sonarsource.scanner.api.internal.shaded.okhttp.internal.connection.RealConnection.connectSocket(RealConnection.java:265)
    at org.sonarsource.scanner.api.internal.shaded.okhttp.internal.connection.RealConnection.connect(RealConnection.java:183)
    at org.sonarsource.scanner.api.internal.shaded.okhttp.internal.connection.ExchangeFinder.findConnection(ExchangeFinder.java:224)
    at org.sonarsource.scanner.api.internal.shaded.okhttp.internal.connection.ExchangeFinder.findHealthyConnection(ExchangeFinder.java:108)
    at org.sonarsource.scanner.api.internal.shaded.okhttp.internal.connection.ExchangeFinder.find(ExchangeFinder.java:88)
    at org.sonarsource.scanner.api.internal.shaded.okhttp.internal.connection.Transmitter.newExchange(Transmitter.java:169)
    at org.sonarsource.scanner.api.internal.shaded.okhttp.internal.connection.ConnectInterceptor.intercept(ConnectInterceptor.java:41)
```

Comando para ejecutar SonarScanner

```
C:\Testing\sonarqube-10.2.1.78527\bin\windows-x86-64>StartSonar.bat
Starting SonarQube...
2024.01.18 10:05:42 INFO app[[o.s.a.AppFileSystem] Cleaning or creating temp directory C:\Testing\sonarqube-10.2.1.78527\temp
2024.01.18 10:05:42 INFO app[[o.s.a.es.EsSettings] Elasticsearch listening on [HTTP: 127.0.0.1:9001, TCP: 127.0.0.1:62249]
2024.01.18 10:05:42 INFO app[[o.s.a.ProcessLauncherImpl] Launch process[ELASTICSEARCH] from [C:\Testing\sonarqube-10.2.1.78527\elasticsearch]: C:\Program
Files\Java\jdk-17\bin\java -Xms4m -Xmx64m -XX:+UseSerialGC -Dcli.name=server -Dcli.script=/bin/elasticsearch -Dcli.libs=lib/tools/server-cli -Des.path.home
=C:\Testing\sonarqube-10.2.1.78527\elasticsearch -Des.path.conf=C:\Testing\sonarqube-10.2.1.78527\temp\conf\es -Des.distribution.type=tar -cp C:\Testing\son
arqube-10.2.1.78527\elasticsearch\lib\*;C:\Testing\sonarqube-10.2.1.78527\elasticsearch\lib\cli-launcher\* org.elasticsearch.launcher.CliToolLauncher
2024.01.18 10:05:42 INFO app[[o.s.a.SchedulerImpl] Waiting for Elasticsearch to be up and running
2024.01.18 10:05:43 INFO app[[o.s.a.SchedulerImpl] Process[es] is up
2024.01.18 10:05:43 INFO app[[o.s.a.ProcessLauncherImpl] Launch process[WEB_SERVER] from [C:\Testing\sonarqube-10.2.1.78527]: C:\Program Files\Java\jdk-17
\bin\java -Djava.awt.headless=true -Dfile.encoding=UTF-8 -Djava.io.tmpdir=C:\Testing\sonarqube-10.2.1.78527\temp -XX:-OmitStackTraceInFastThrow --add-opens=
java.base/java.util=ALL-UNNAMED --add-opens=java.base/java.lang=ALL-UNNAMED --add-opens=java.base/java.io=ALL-UNNAMED --add-opens=java.rmi/sun.rmi.transp
ort=ALL-UNNAMED --add-exports=java.base/jdk.internal.ref=ALL-UNNAMED --add-opens=java.base/java.nio=ALL-UNNAMED --add-opens=java.nio.ch=ALL-UNNAMED --
add-opens=java.management/sun.management=ALL-UNNAMED --add-opens=jdk.management/com.sun.management.internal=ALL-UNNAMED -Dcom.redhat.fips=false -Xmx512m -Xm
s128m -XX:+HeapDumpOnOutOfMemoryError -Dhttp.nonProxyHosts=localhost[127.*][:1] -cp ./lib/sonar-application-10.2.1.78527.jar;C:\Testing\sonarqube-10.2.1.78
527\lib\jdbc\h2\h2-2.1.214.jar org.sonar.server.app.WebServer C:\Testing\sonarqube-10.2.1.78527\temp\sq-process1176047338217788247properties
WARNING: A terminally deprecated method in java.lang.System has been called
WARNING: System:setSecurityManager has been called by org.sonar.process.PluginSecurityManager (file:/C:/Testing/sonarqube-10.2.1.78527/lib/sonar-applicatio
n-10.2.1.78527.jar)
WARNING: Please consider reporting this to the maintainers of org.sonar.process.PluginSecurityManager
WARNING: System:setSecurityManager will be removed in a future release
2024.01.18 10:05:43 INFO app[[o.s.a.SchedulerImpl] Process[Web Server] is stopped
2024.01.18 10:05:43 INFO app[[o.s.a.SchedulerImpl] Process[ElasticSearch] is stopped
2024.01.18 10:05:43 INFO app[[o.s.a.SchedulerImpl] SonarQube is stopped

C:\Testing\sonarqube-10.2.1.78527\bin\windows-x86-64>
```

Nota: el servicio levanta el localhost:900

Anexo 6. Resultados a nivel inferencial

En este apartado se indica los resultados según las dimensiones del trabajo investigación, dichas dimensiones son: escalabilidad, seguridad y rendimiento. Para la obtención de los resultados de la mejora de escalabilidad, seguridad y rendimiento, los cuales fueron hechos antes y después del uso de ChatGPT.

Prueba de la hipótesis general

Se relata la información estadística inferencial, asimismo se obtiene los datos de la variable dependiente, el cual se realizó a nuestra unidad de análisis mediante el instrumento con el fin de medir la mejora de la calidad del desarrollo web por el uso del ChatGPT.

Tabla 1 Mejora de la calidad del desarrollo web

		Estadístico	Error estándar
PRETEST_VD	Media	45,40	3,0703
POSTEST_VD	Media	22,80	3,1581

La tabla detalla la media del pre test y post test, teniendo 45,40 en la prueba inicial y 22,80 en la final. La media de cada test sirve para determinar el incremento porcentual de la calidad del desarrollo web.

Tabla 2 Prueba de normalidad de mejora de calidad del desarrollo web

Prueba Shapiro-Wilk				
	Estadístico	gl	Sig.	
PRETEST_VD	0,848	10	0,055	
POSTEST_VD	0.812	10	0,020	

Para la variable dependiente se empleó la prueba Shapiro-Wilk, porque tiene una muestra menor a 50. La tabla indica que los datos del resultado estadístico de normalidad se obtuvieron al realizar la prueba, los cuales son 0.848 y 0.812 referente a la prueba inicial y final respectivamente.

En el pretest se obtuvo un dato estadístico de 0.848 con una significancia por encima de 0.05, otorgándonos una distribución normal.

En el posttest se obtuvo un dato estadístico de 0.812 con una significancia por debajo de 0.05, otorgándonos una distribución no normal.

Hipótesis específica (HG)

H0: El uso de ChatGPT tiene un impacto negativo en la calidad del desarrollo web en la empresa Investigación S.A.C.

H1: El uso de ChatGPT tiene un impacto positivo en la calidad del desarrollo web en la empresa Investigación S.A.C.

Prueba de Wilcoxon

Se procede a realizar la prueba de Wilcoxon con la finalidad de hacer las comparaciones de ambos valores expresados a continuación:

Tabla 3 Rango con signo de Wilcoxon - Mejora de calidad del desarrollo web

		N	Rango promedio	Suma de rangos
POSTEST_VD - PRETEST_VD	R. negativos	10a	5,50	55,00
	R. positivos	0b	0,00	0,00
	Empates	0c		
	Total	10		

a) POSTEST_VD < PRETEST_VD

b) POSTEST_VD > PRETEST_VD

c) POSTEST_VD = PRETEST_VD

La tabla señala que 10 son rangos negativos, su promedio es 5.50 y la suma es 55, no teniendo rangos positivos ni empates en la muestra.

Tabla 4 Estadístico de prueba Wilcoxon - Mejora de calidad del desarrollo web

	POSTEST_VD - PRETEST_VD
Z	-2,805
Sig. asintótica(bilateral)	0,005

La tabla muestra los datos estadísticos del test con un valor de $Z = -2.805$ donde Z pertenece a la zona de rechazo. También el valor de P es menor de 0.05, por tal razón se está aceptando la hipótesis alterna con un nivel de confianza al 95%. Por ello se indica que el uso de ChatGPT tiene un impacto positivo en la calidad del desarrollo web en la empresa Investigación S.A.C.

Para obtener incremento de porcentaje de la calidad del desarrollo web se analiza las medias de cada test, teniendo en la prueba inicial una media de 22,80, y en la final una media de 45,40, la cual señala un aumento del 99,12%. Se adjunta, el análisis operacional para obtener el porcentaje:

IVD = Incremento de variable dependiente

POSTEST_VD = Variable dependiente pos-prueba

PRETEST_VD = Variable dependiente pre-prueba

Prueba de la hipótesis específica 1

Se relata la información estadística inferencial, asimismo se obtiene los datos de la dimensión escalabilidad, el cual se realizó a nuestra unidad de análisis mediante el instrumento con el fin de medir la mejora de la escalabilidad en la calidad del desarrollo web por el uso del ChatGPT.

Tabla 5 Mejora de la escalabilidad

		Estadístico	Error estándar
PRETEST_DIM01	Media	793,30	11,9926
POSTEST_DIM01	Media	148,30	17,5348

La tabla detalla la media del pre test y post test, teniendo 793,30 en la prueba inicial y 148,30 en la final. La media de cada test sirve para determinar el incremento porcentual de la escalabilidad.

Tabla 6 Prueba de normalidad de mejora de la escalabilidad

	Prueba Shapiro-Wilk			
	Estadístico	gl	Sig.	
PRETEST_DIM01	0,922	10	0,370	
POSTEST_DIM01	0.709	10	0,001	

Para la dimensión 1 se empleó la prueba Shapiro-Wilk, porque tiene una muestra menor a 50. La tabla indica que los datos del resultado estadístico de normalidad se obtuvieron al realizar la prueba, los cuales son 0.922 y 0.709 referente a la prueba inicial y final respectivamente.

En el pretest se obtuvo un dato estadístico de 0.922 con una significancia por encima de 0.05, otorgándonos una distribución normal.

En el postest se obtuvo un dato estadístico de 0.709 con una significancia por debajo de 0.05, otorgándonos una distribución no normal.

Hipótesis específica (HE1)

H0: El uso de ChatGPT en el desarrollo de una web en la empresa Investigación S.A.C 2023 tendrá un efecto negativo en la escalabilidad de la web.

H1: El uso de ChatGPT en el desarrollo de una web en la empresa Investigación S.A.C 2023 tendrá un efecto positivo en la escalabilidad de la web.

Prueba de Wilcoxon

Se procede a realizar la prueba de Wilcoxon con la finalidad de hacer las comparaciones de ambos valores expresados a continuación:

Tabla 7 Rango con signo de Wilcoxon - Mejora de escalabilidad

		N	Rango promedio	Suma de rangos
POSTEST_DIM01- PRETEST_DIM01	R. negativos	10a	5,50	55,00
	R. positivos	0b	0,00	0,00
	Empates	0c		
	Total	10		

a) POSTEST_DIM01 < PRETEST_DIM01

b) POSTEST_DIM01 > PRETEST_DIM01

c) POSTEST_DIM01 = PRETEST_DIM01

La tabla señala que 10 son rangos negativos, su promedio es 5.50 y la suma es 55, no teniendo rangos positivos ni empates en la muestra.

Tabla 8 Estadístico de prueba Wilcoxon - Mejora de escalabilidad

	POSTEST_DIM01- PRETEST_DIM01
Z	-2,805
Sig. asintótica(bilateral)	0,005

La tabla muestra los datos estadísticos del test con un valor de $Z = -2.805$ donde Z pertenece a la zona de rechazo. También el valor de P es menor de 0.05 , por tal razón se está aceptando la hipótesis alterna con un nivel de confianza al 95% . Por ello se indica que el uso de ChatGPT en el desarrollo de una web en la empresa Investigación S.A.C 2023 tendrá un efecto positivo en la escalabilidad de la web.

Para obtener incremento de porcentaje de la escalabilidad se analiza las medias de cada test, teniendo en la prueba inicial una media de $148,30$, y en la final una media de $793,30$, la cual señala un aumento del $434,93\%$. Se adjunta, el análisis operacional para obtener el porcentaje:

IDIM01 = Incremento de variable dependiente

POSTEST_DIM01= Dimensión 1 pos-prueba

PRETEST_DIM01= Dimensión 1 pre-prueba

Prueba de la hipótesis específica 2

Se relata la información estadística inferencial, asimismo se obtiene los datos de la dimensión seguridad, el cual se realizó a nuestra unidad de análisis mediante el instrumento con el fin de medir la mejora de la seguridad en la calidad del desarrollo web por el uso del ChatGPT.

Tabla 9 Mejora de la seguridad

		Estadístico	Error estándar
PRETEST_DIM02	Media	10,80	2,3181
POSTEST_DIM02	Media	6,00	1,5564

La tabla detalla la media del pre test y post test, teniendo 10,80 en la prueba inicial y 6,00 en la final. La media de cada test sirve para determinar el incremento porcentual de la seguridad.

Tabla 10 Prueba de normalidad de mejora de la seguridad

Prueba Shapiro-Wilk				
	Estadístico	gl	Sig.	
PRETEST_DIM02	0,783	10	0,009	
POSTEST_DIM02	0.837	10	0,041	

Para la dimensión 2 se empleó la prueba Shapiro-Wilk, porque tiene una muestra menor a 50. La tabla indica que los datos del resultado estadístico de normalidad se obtuvieron al realizar la prueba, los cuales son 0.783 y 0.837 referente a la prueba inicial y final respectivamente.

En el pretest se obtuvo un dato estadístico de 0.783 con una significancia por debajo de 0.05, otorgándonos una distribución no normal.

En el posttest se obtuvo un dato estadístico de 0.837 con una significancia por encima de 0.05, otorgándonos una distribución normal.

Hipótesis específica (HE2)

H0: El uso de ChatGPT en el desarrollo de una web en la empresa Investigación S.A.C 2023 tendrá un efecto negativo en la seguridad de la web.

H1: El uso de ChatGPT en el desarrollo de una web en la empresa Investigación S.A.C 2023 tendrá un efecto positivo en la seguridad de la web.

Prueba de Wilcoxon

Se procede a realizar la prueba de Wilcoxon con la finalidad de hacer las comparaciones de ambos valores expresados a continuación:

Tabla 11 Rango con signo de Wilcoxon - Mejora de seguridad

		N	Rango promedio	Suma de rangos
POSTEST_DIM02- PRETEST_DIM02	R. negativos	10a	5,69	45,50
	R. positivos	0b	4,75	9,50
	Empates	0c		
	Total	10		

a) POSTEST_DIM02 < PRETEST_DIM02

b) POSTEST_DIM02 > PRETEST_DIM02

c) POSTEST_DIM02 = PRETEST_DIM02

La tabla señala que 8 son rangos negativos, su promedio es 5.69 y la suma es 45,50, además 2 son rangos positivos, su promedio es 4.75 y la suma es 9,50, no teniendo rangos empates en la muestra.

Tabla 12 Estadístico de prueba Wilcoxon - Mejora de seguridad

	POSTEST_DIM02- PRETEST_DIM02
Z	-1,838
Sig. asintótica(bilateral)	0,007

La tabla muestra los datos estadísticos del test con un valor de $Z = -1.838$ donde Z pertenece a la zona de rechazo. También el valor de P es menor de 0.05 , por tal razón se está aceptando la hipótesis alterna con un nivel de confianza al 95% . Por ello se indica que el uso de ChatGPT en el desarrollo de una web en la empresa Investigación S.A.C 2023 tendrá un efecto positivo en la seguridad de la web.

Para obtener incremento de porcentaje de la seguridad se analiza las medias de cada test, teniendo en la prueba inicial una media de $10,80$, y en la final una media de $6,00$, la cual señala un aumento del $80,00\%$. Se adjunta, el análisis operacional para obtener el porcentaje:

IDIM02 = Incremento de variable dependiente

POSTEST_DIM02= Dimensión 2 pos-prueba

PRETEST_DIM02= Dimensión 2 pre-prueba

Prueba de la hipótesis específica 3

Se relata la información estadística inferencial, asimismo se obtiene los datos de la dimensión rendimiento, el cual se realizó a nuestra unidad de análisis mediante el instrumento con el fin de medir la mejora del rendimiento en la calidad del desarrollo web por el uso del ChatGPT.

Tabla 13 Mejora del rendimiento

		Estadístico	Error estándar
PRETEST_DIM03	Media	14,60	2,4549
POSTEST_DIM03	Media	14,80	3,9237

La tabla detalla la media del pre test y post test, teniendo 10,80 en la prueba inicial y 6,00 en la final. La media de cada test sirve para determinar el incremento porcentual del rendimiento.

Tabla 14 Prueba de normalidad de mejora del rendimiento

Prueba Shapiro-Wilk				
	Estadístico	gl	Sig.	
PRETEST_DIM03	0,946	10	0,623	
POSTEST_DIM03	0.804	10	0,016	

Para la dimensión 3 se empleó la prueba Shapiro-Wilk, porque tiene una muestra menor a 50. La tabla indica que los datos del resultado estadístico de normalidad se obtuvieron al realizar la prueba, los cuales son 0.946 y 0.804 referente a la prueba inicial y final respectivamente.

En el pretest se obtuvo un dato estadístico de 0.946 con una significancia por encima de 0.05, otorgándonos una distribución normal.

En el postest se obtuvo un dato estadístico de 0.804 con una significancia por debajo de 0.05, otorgándonos una distribución no normal.

Hipótesis específica (HE3)

H0: El uso de ChatGPT en el desarrollo de una web en la empresa Investigación S.A.C 2023 tendrá un efecto negativo en el rendimiento de la web.

H1: El uso de ChatGPT en el desarrollo de una web en la empresa Investigación S.A.C 2023 tendrá un efecto positivo en el rendimiento de la web.

Prueba de Wilcoxon

Se procede a realizar la prueba de Wilcoxon con la finalidad de hacer las comparaciones de ambos valores expresados a continuación:

Tabla 15 Rango con signo de Wilcoxon - Mejora de rendimiento

		N	Rango promedio	Suma de rangos
POSTEST_DIM02- PRETEST_DIM02	R. negativos	10a	5,50	55,00
	R. positivos	0b	0,00	0,00
	Empates	0c		
	Total	10		

a) POSTEST_DIM03 < PRETEST_DIM03

b) POSTEST_DIM03 > PRETEST_DIM03

c) POSTEST_DIM03 = PRETEST_DIM03

La tabla señala que 10 son rangos negativos, su promedio es 5.50 y la suma es 55,00, no teniendo rangos positivos ni empates en la muestra.

Tabla 16 Estadístico de prueba Wilcoxon - Mejora de rendimiento

	POSTEST_DIM03- PRETEST_DIM03
Z	-1,780
Sig. asintótica(bilateral)	0,009

La tabla muestra los datos estadísticos del test con un valor de $Z = -1.780$ donde Z pertenece a la zona de rechazo. También el valor de P es menor de 0.05 , por tal razón se está aceptando la hipótesis alterna con un nivel de confianza al 95% . Por ello se indica que el uso de ChatGPT en el desarrollo de una web en la empresa Investigación S.A.C 2023 tendrá un efecto positivo en el rendimiento de la web.

Para obtener incremento de porcentaje del rendimiento se analiza las medias de cada test, teniendo en la prueba inicial una media de $14,60$, y en la final una media de $14,80$, la cual señala un aumento del $1,37\%$. Se adjunta, el análisis operacional para obtener el porcentaje:

IDIM03 = Incremento de variable dependiente

POSTEST_DIM03= Dimensión 3 pos-prueba

PRETEST_DIM03= Dimensión 3 pre-prueba



ESCUELA DE POSGRADO

MAESTRÍA EN INGENIERÍA DE SISTEMAS CON MENCIÓN EN TECNOLOGÍAS DE LA INFORMACIÓN

Declaratoria de Autenticidad del Asesor

Yo, ALZA SALVATIERRA SILVIA DEL PILAR, docente de la ESCUELA DE POSGRADO MAESTRÍA EN INGENIERÍA DE SISTEMAS CON MENCIÓN EN TECNOLOGÍAS DE LA INFORMACIÓN de la UNIVERSIDAD CÉSAR VALLEJO SAC - LIMA NORTE, asesor de Tesis titulada: "Uso de chatGPT en la calidad del desarrollo web de una empresa privada de Lima, 2023", cuyo autor es ROBLES LUJAN DARWIN EDWIN, constato que la investigación tiene un índice de similitud de 12.00%, verificable en el reporte de originalidad del programa Turnitin, el cual ha sido realizado sin filtros, ni exclusiones.

He revisado dicho reporte y concluyo que cada una de las coincidencias detectadas no constituyen plagio. A mi leal saber y entender la Tesis cumple con todas las normas para el uso de citas y referencias establecidas por la Universidad César Vallejo.

En tal sentido, asumo la responsabilidad que corresponda ante cualquier falsedad, ocultamiento u omisión tanto de los documentos como de información aportada, por lo cual me someto a lo dispuesto en las normas académicas vigentes de la Universidad César Vallejo.

LIMA, 17 de Enero del 2024

Apellidos y Nombres del Asesor:	Firma
ALZA SALVATIERRA SILVIA DEL PILAR DNI: 18110381 ORCID: 0000-0002-7075-6167	Firmado electrónicamente por: SALZAS el 20-01- 2024 09:42:43

Código documento Trilce: TRI - 0733911