

A systematic review of the effectiveness of the use of open source software in web observatories

César Herrera-Lozano, estudiante ingeniería de sistemas e informática¹, Juan Vásquez-Preciado, estudiante ingeniería de sistemas e informática², Christian Dios-Castillo, Dr en Administración de la Educación³, Gilberto Carrión-Barco, Dr en Ciencias de la Computación y Sistemas⁴

^{1,2,3,4}Universidad Tecnológica del Perú, Perú, U19200293@utp.edu.pe, 1601033@utp.edu.pe, cdios@utp.edu.pe, C14051@utp.edu.pe

Abstract– Since much of the Internet is now based on free and open source software (FOSS), open source is becoming the norm in software development, FOSS has been so successful that for many applications it is the de facto standard, with 94% of the world's top 500 supercomputers, 75% of the top 10,000 web sites, and 98% of companies using open source software. The purpose of this study is to evaluate the effectiveness of implementing open source solutions in web observatories, identifying which web observatories use open source solutions in order to measure that the use can be effective and of quality. The present study was conducted following the guidelines of an SLR, which consists of three PICO questions, to select relevant publications, inclusion and exclusion criteria were applied, and articles were retrieved from the Scopus database following the PRISMA procedure. Different advantages of the use of CA at the time of its implementation were found that make its practice effective. Likewise, success stories were identified in which CA and methods were used to measure the effectiveness and quality of a web observatory. In conclusion, this study demonstrates that the use of open source is a viable option with various applications because it is free and avoids incurring software licensing costs and other associated expenses.

Keywords-- open source, web observatory, efficiency, productivity.

Una revisión sistemática sobre la efectividad del uso de software de código abierto en observatorios web

César Herrera-Lozano, estudiante ingeniería de sistemas e informática¹, Juan Vásquez-Preciado, estudiante ingeniería de sistemas e informática², Christian Dios-Castillo, Dr en Administración de la Educación³, Gilberto Carrión-Barco, Dr en Ciencias de la Computación y Sistemas⁴

^{1,2,3,4}Universidad Tecnológica del Perú, Perú, U19200293@utp.edu.pe, 1601033@utp.edu.pe, cdios@utp.edu.pe, C14051@utp.edu.pe

Resumen- Dado que gran parte de Internet se basa ahora en software libre y de código abierto (FOSS), el código abierto se está convirtiendo en la norma en el desarrollo de software, el FOSS ha tenido tanto éxito que para muchas aplicaciones es el estándar de facto, con el 94% de los 500 superordenadores más importantes del mundo, el 75% de los 10.000 sitios web más importantes y el 98% de las empresas que utilizan software de código abierto. El objetivo de este estudio es evaluar la eficacia de la implementación de soluciones de código abierto en los observatorios web, identificando qué observatorios web utilizan soluciones de código abierto para medir que el uso puede ser eficaz y de calidad. El presente estudio se realizó siguiendo las directrices de un SLR, que consta de tres preguntas PICO, para seleccionar las publicaciones relevantes, se aplicaron criterios de inclusión y exclusión, y se recuperaron artículos de la base de datos Scopus siguiendo el procedimiento PRISMA. Se encontraron diferentes ventajas del uso de la AC en el momento de su implantación que hacen efectiva su práctica. Asimismo, se identificaron casos de éxito en los que se utilizaron AC y métodos para medir la eficacia y la calidad de un observatorio web. En conclusión, este estudio demuestra que el uso de código abierto es una opción viable con diversas aplicaciones, ya que es gratuito y evita incurrir en costes de licencias de software y otros gastos asociados.

Palabras clave-- código abierto, observatorio web, eficacia, productividad.

I. INTRODUCCIÓN

La innovación tecnológica exige el uso adecuado de tecnologías digitales, y la creatividad no se expresa solamente en la creación de código de algoritmos, sino en la integración de estos, siendo de nuestra creación o de acceso libre. Dado que gran parte de internet ahora se basa en software libre y de código abierto (FOSS), el código abierto se está convirtiendo en la norma en el desarrollo de software, FOSS ha tenido tanto éxito que para muchas aplicaciones es el estándar de facto, con el 94% de las 500 supercomputadoras más destacadas a nivel mundial, el 75% de los 10 000 sitios web principales y el 98% de las empresas que utilizan software de código abierto [1]. Asimismo, el motor de búsqueda posibilita la obtención de documentos disponibles en la web provenientes de editoriales académicas, sociedades científicas y profesionales, repositorios digitales, universidades y otros sitios web que cumplen con una serie de requisitos técnicos [2]. Si bien el modelo de CA permite que muchas personas revisen y auditen el código, también puede exponer a los proyectos a un mayor escrutinio y riesgo de vulnerabilidades. La identificación y corrección de fallas de seguridad pueden requerir un esfuerzo

colaborativo, y algunas organizaciones pueden ser menos diligentes en la aplicación de “parches” y actualizaciones, también se realizó una conceptualización sobre el significado de software libre y la naturaleza de las aplicaciones seleccionadas para el estudio, con el propósito de recopilar los datos de los sistemas de información objeto de investigación [3]. En Perú, este enfoque colaborativo ha sido reconocido como un mecanismo prometedor para promover la innovación, transparencia para el acceso a la tecnología en diferentes sectores. Se han realizado esfuerzos para fomentar el uso al desarrollar el software de código abierto en el gobierno, la educación e industria privada, la implementación de una plataforma de aprendizaje, el Open Source en Perú continúa evolucionando y desempeña un papel fundamental en la promoción de una cultura tecnológica abierta y colaborativa en el país [4]. El software de código abierto (OS) ofrece beneficios superiores en comparación con el software adquirido debido a los derechos que otorga a sus usuarios. Estas ventajas son valoradas tanto por usuarios individuales como por empresas y entidades gubernamentales. Además, bibliotecas e instituciones de América Latina, como Colombia, Chile y Perú, han llevado a cabo proyectos de automatización utilizando software libre, obteniendo resultados positivos de manera fortuita [5]. La presente investigación nos brinda información valiosa para futuros estudios sobre el uso de software libre en observatorios web, lo que puede ser útil para la comunidad de ingeniería de sistemas e informática. Además, nos presentan artículos que pueden ser utilizados como base para futuras investigaciones en el área. En resumen, este artículo puede ser una fuente de información importante para aquellos interesados en el uso de software libre en observatorios web. El objetivo de esta investigación es determinar la efectividad del uso de soluciones de código abierto (OPEN SOURCE) en observatorios web, identificando que observatorios web utilizan soluciones de código abierto para así medir que el uso puede ser efectivo y de calidad. Se menciona que el desarrollo tecnológico de código abierto es un sistema descentralizado, participativo y transparente, en el cual numerosos voluntarios con diversos conocimientos contribuyen de manera autónoma al desarrollo. Buscando evaluar cómo el uso de software de código abierto puede beneficiar a los observatorios web en términos de su funcionalidad, eficiencia y capacidad para recopilar y analizar datos de manera efectiva. Asimismo, implica analizar las ventajas y desventajas de utilizar OS en comparación con

soluciones propietarias, se puede inferir que el artículo sugiere que el uso de soluciones de código abierto es una opción viable y beneficiosa para los observatorios web. En opinión de los autores (Rebecca C. Jackson1 James P. Balhoff2, Eric Douglass3 Nomi L. Harris3, Christopher J. Mungall3, James A. Overton1) podemos visualizar que adoptan un enfoque descriptivo para así implementar la última generación de Código abierto para facilitar la investigación sobre aplicaciones de datos de alta dimensión, con el fin de brindar una función valiosa para que así se pueda analizar datos en software libre [6]. Teniendo en cuenta que el open source es un desafío clave para obtener en la visualización eficiente de datos [7]. Lo más conspirador es que se demuestre que el OS puede funcionar incluso cuando los investigadores no estén realizando este tipo de análisis en sus datos [8] El OS juega un papel importante en la ingeniería de software moderna, con el aumento de la cantidad de código fuente disponible en internet, ayudan a los desarrolladores de software a recuperar y reutilizar el código existente con mayor eficacia .Se necesita idear formas nuevas y fáciles para proporcionar información auxiliar al proceso de búsqueda para complementar las especificaciones débiles, para poder generar nuevas entradas y salidas para los métodos válidos basados en el análisis dinámico y proporcionar a los usuarios la comportamientos adicionales para ayudarlos a identificar los métodos correctos fácilmente [9]. El uso de OS que es estándar en la comunidad de investigación permite que los usuarios puedan acceder y ampliar el código de programación, la necesidad de hacer que el software sea de código abierto permitiendo que el código de programación esté disponible para todos y que los futuros usuarios con experiencia en programación puedan ampliar las opciones analíticas [10]. Finalmente, las licencias de software abierto son tanto un modelo de desarrollo como un instrumento legal, el desarrollo abierto implica que los proyectos construyen comunidades de forma activa mediante el uso de plataformas de código compartido, canales de redes sociales, talleres periódicos y otras formas de participación. El OS debe verse como una aspiración: no es una condición necesaria para la transparencia pública o la reproducibilidad científica [11][12].

II. METODOLOGÍA

A. Estrategia de búsqueda

El presente estudio se realizó siguiendo las pautas de una revisión sistemática de literatura, esta metodología tiene como propósito identificar, examinar e interpretar exhaustivamente toda la evidencia pertinente asociadas con una pregunta de investigación determinada [13]. Para ello se estableció la siguiente pregunta PICO: la efectividad del open source para observatorios web, esta misma se descompone en nuevas preguntas las cuales se plasman en la siguiente tabla.

TABLA I
ESTRUCTURA PICO

PICO	Pregunta	Palabras clave
------	----------	----------------

P	¿Cuál es la efectividad del uso de soluciones de código abierto (Open Source) en observatorios web?	Open source software, open source
I	¿Qué Observatorios web que utilizan soluciones de código abierto?	web observatory, web-based observatory, web application
C	-	-
O	¿Cómo se mide la efectividad y calidad de los webs sites?	efficiency, productivity, effectiveness

Este estudio está basado en una búsqueda de literatura utilizando como única base de datos a SCOPUS, lo que implica buscar términos de forma automatizada.

La fórmula de búsqueda utilizada en este estudio fue:

(TITLE-ABS-KEY (“efficiency” OR “productivity” OR “effectiveness”) AND TITLE-ABS-KEY (“web sites” OR “web*” OR “web application” OR “web observatory” OR “web-based observatory” OR “web-based monitoring”) AND TITLE-ABS-KEY (“open source” OR “open source software”))

B. Criterios de inclusión y exclusión

En base a los objetivos del presente estudio se tomaron en cuenta 3 criterios de inclusión mostrados en la siguiente tabla:

TABLA II
CRITERIOS DE INCLUSIÓN

CI1: Los estudios incluidos deben abordar el uso de open source
CI2: Los estudios incluidos deben estar relacionados con la efectividad del open source
CI3: Los estudios incluidos deben estar relacionados a observatorios web o sitios web.

Asimismo, los criterios de exclusión que se tomaron en cuenta para el presente estudio responden a los 8 criterios mostrados en la siguiente tabla:

TABLA III
CRITERIOS DE EXCLUSIÓN

CE1: Publicaciones anteriores a 2013 y posteriores a 2022
CE2: Tipo de publicación no corresponde a artículo
CE3: Artículos sin acceso abierto
CE4: Publicaciones diferentes al idioma inglés y español
CE5: Publicaciones con menos de 3 citas
CE6: Artículos con título y resumen ajenos al tema de investigación
CE7: Artículos a texto completo ajenos al tema de investigación.

Fig 1. Flujo de trabajo de planificación, identificación y elegibilidad.

C. Proceso de selección de estudios

Luego de identificar 1278 artículos obtenidos en SCOPUS, para este estudio se ha considerado adecuado realizar una revisión sistemática basada en la declaración PRISMA.

La declaración PRISMA es beneficioso al momento de organizar y llevar a cabo revisiones sistemáticas para asegurarse de recopilar toda la información sugerida, permite al lector analizar la idoneidad de los métodos y, por lo tanto, la credibilidad de los hallazgos [14]. Consiste en obtener resultados de publicaciones científicas relacionadas al objetivo de la investigación, este protocolo se basa en dos etapas de filtrado en la primera etapa, se filtra según los criterios de inclusión y exclusión establecidos para la investigación en base a la revisión de título, resumen, año de publicación, etc. En la segunda etapa, se hace el filtrado a partir de los criterios de inclusión y exclusión establecidos para la investigación en base a una revisión a texto completo.

A partir de 1278 artículos obtenidos, se inició la primera fase de filtrado en base a los criterios de inclusión y exclusión mencionados anteriormente. Aplicando estos criterios, se descartaron n=1213 publicaciones por las siguientes razones: n=381 no fueron publicados entre los años 2013 – 2022, n=534 no fueron publicaciones de tipo artículo, n=162 por no ser publicaciones de acceso abierto, n=3 por ser publicaciones en idiomas diferente al inglés y español, n=61 por ser publicaciones con menos de tres citas. Posteriormente, y tras la lectura del título y resumen, se eliminaron n=67 tras no tener relación con nuestro estudio. Asimismo, n=5 publicaciones no se pudieron recuperar a texto completo (PDF o HTML). Esto dejó un total de n=65 registros seleccionados para evaluar su elegibilidad. Finalmente, después de leer los artículos seleccionados en su totalidad, se descartaron n=17 por ser ajenos al tema de investigación y n=13 por estar enfocados en otra línea de ingeniería. Por lo tanto, se incluyeron n=35 registros para la revisión sistemática.

III. RESULTADOS

A continuación, se muestran los artículos estudiados en la presente investigación, los cuales están organizados por características como año de publicación, país y área temática. Al analizar la figura 2 correspondiente a los artículos organizados por años se pudo verificar que la producción científica tuvo un aumento en el 2019, encontrándose 10 de los 35 artículos incluidos para el análisis.

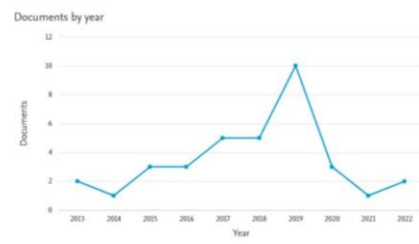


Figura 2. Artículos organizados por fecha de publicación.

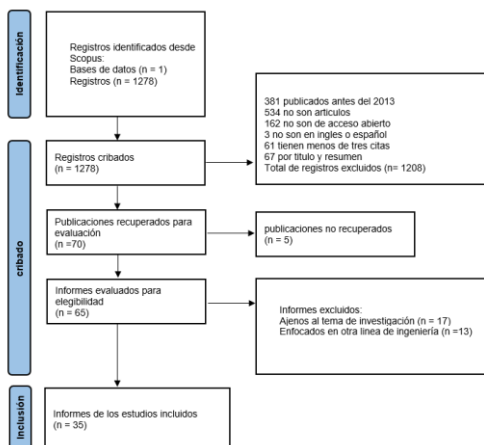
En referencia al análisis de los artículos organizados por países, 8 de 35 artículos incluidos para esta investigación fueron publicados en Estados Unidos, tal como se muestra en la figura 3.



Figura 3. Artículos organizados por país de publicación.

Fuente: Autores

Asimismo, al analizar la figura 4 de artículos organizados por área temática, se observa que la mayoría pertenece al área de ciencias de la computación (22.7%), seguidos de los estudios de ingeniería (14.8%).



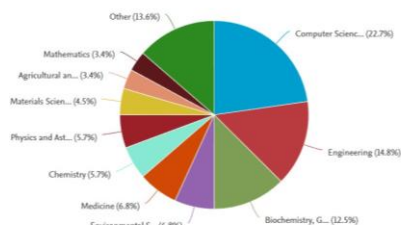


Figura 4. Artículos organizados por área temática.

A continuación, se realiza una descripción cualitativa de los estudios seleccionados para el presente estudio, en función de las preguntas de investigación planteadas al principio.

a. ¿Cuál es la efectividad del uso de soluciones de código abierto (Open Source) en observatorios web?

El desarrollo tecnológico de CA y gratuito es un sistema completamente nuevo que se caracteriza por ser descentralizado, participativo y transparente [1], es un enfoque de diseño que ha surgido, en la cual numerosos voluntarios con diversos conocimientos contribuyen de manera autónoma al desarrollo [15]. Al emplear soluciones de CA trae consigo ventajas de implementación que hacen de este ser efectivo. K. S. Burgdorf utiliza un software de CA debido a que puede adquirirlo a bajo costo o costo cero [16]. Asimismo, la capacidad de aprovechar bibliotecas y software de CA es una característica destacada, ya que elimina la necesidad de incurrir en costos de licencias de software y otros gastos asociados [17]. Al combinar las soluciones de CA y sus costos hacen que requiriera menos capital de inversión inicial [18], lo que hace efectivo su uso en cuestiones de dinero. Además, en cuestiones de tiempo, se ha observado que la utilización de software de CA que son ágiles, económicas y adaptables para la configuración, acelera el proceso de desarrollo de software en cuestión de semanas [19], al posibilitar a los usuarios reutilizar, intercambiar o mejorar módulos individuales sin que ello tenga repercusiones en otros componentes del sistema [20], [21]. Para los desarrolladores que deseen experimentar con este nuevo enfoque de trabajo, representa una oportunidad valiosa sin la necesidad de realizar inversiones significativas o asumir compromisos que consuman mucho tiempo [22]. El uso de tecnologías de CA para la creación de una plataforma no restringe el rendimiento global del sistema [23], teniendo resultados similares o mejores en comparación con las tecnologías que requieren de licencias. Asimismo, al ser tecnologías de CA, se previenen infracciones de derechos de autor [24]. La colaboración y comunidad activa son razones que hacen del CA ser efectivo, en GitHub, la amplia variedad de proyectos de CA ofrece a individuos de diferentes partes del mundo la oportunidad de colaborar sin necesidad de reunirse físicamente. Además, los desarrolladores gozan de mayor libertad en su trabajo. Asimismo, la multitud de proyectos de CA alojados públicamente proporciona a los desarrolladores oportunidades para contribuir, ampliando los grupos de colaboradores más allá de un reducido núcleo [12].

La escalabilidad y flexibilidad de soluciones de CA contribuyen a su uso efectivo, Yang, Chen, Chan y Shen desarrollan e implementa un servicio de almacenamiento en la nube mediante la integración de software de CA, con la finalidad de brindar un servicio de almacenamiento a los usuarios para facilitar un desarrollo y mantenimiento más efectivo a largo plazo [25]. Esto significa que el software puede adaptarse a medida que crece y evoluciona permitiendo mayor interoperabilidad y expansión de funcionalidades. La transparencia y personalización del CA hace que su uso se vuelva efectivo. Conforme el uso de soluciones de código abierto se vuelve más frecuente en el desarrollo de nuevos sistemas de software, el fácil acceso a este hace posible buscar y utilizar fragmentos de código reutilizables, ya sea tal cual o con modificaciones, para abordar tareas de programación específicas evitando código repetitivo [26], [27].

TABLA IV
ESTUDIOS SOBRE LAS VENTAJAS DE USAR CA

Descripción	Documentos
Sobre el código abierto	Wijnen et al [1] Zhang and Li [15]
Ventajas que hacen efectivo el uso de CA	Mao et al [12] Zhang and Li [15] Burgdorf et al [16] Zhang et al [17] Perez-padillo et al [18] Kim et al [19] Yuan et al [20] Wang et al [21] Zaytsev and Morrison [22] Amaxilatis et al [23] A. Ettlin [24] Yang et al [25] Z. Shen [26] A. Fabregat et al [27]

b. ¿Qué Observatorios web que utilizan soluciones de código abierto?

Sólo algunas compañías, principalmente motores de búsqueda web comerciales como Google y Bing, utilizan su propia infraestructura personalizada para la búsqueda. Para el resto de la industria, el motor de búsqueda de código abierto "Lucene" se ha convertido en la opción principal para desarrollar y aplicar en sistemas de búsqueda [28]. MetaboLights es el primer repositorio de acceso abierto y de uso general para estudios de metabolómica. Se divide en dos partes principales: un depósito que permite a la comunidad de metabolómica compartir descubrimientos, datos y protocolos para diversos tipos de estudios de metabolómica, y una capa de conocimiento de referencia que almacena y muestra una amplia gama de información asociada con los estudios en MetaboLights. El repositorio de MetaboLights se basa en software y herramientas de CA que están disponibles de forma gratuita. La aplicación web se despliega en un servidor Apache Tomcat y la información se almacena en una base de datos, subyacente utiliza Oracle como motor, aunque también se pueden utilizar otras bases de datos SQL

estándar como MySQL y PostgreSQL [29]. De la misma forma, Tidepool es una plataforma que ofrece un centro centralizado para los datos de dispositivos de diabetes de los pacientes. Además, proporciona un acceso abierto a aplicaciones de terceros a través de interfaces de programas de aplicaciones (API) RESTful. Utilizar la plataforma Tidepool para recopilar y alojar datos permitiría a los investigadores evitar la necesidad de configurar sistemas de recopilación de datos independientes para cada estudio de investigación, lo que mejoraría la eficiencia y reduciría los costos de investigación. Los productos de software cliente de la plataforma se desarrollan como aplicaciones web utilizando JavaScript y HTML. Al almacenar y distribuir el software a través de la web, los usuarios siempre tienen acceso a la versión más reciente, lo que les permite beneficiarse de forma transparente de los avances y correcciones de errores más recientes [30]. A. M. Fisher, elaboró un proyecto llamado RxMAGIC de código abierto para la gestión de recetas y control de inventario para una clínica, fue desarrollado haciendo uso de tecnologías como Ruby y MySQL [31]. El observatorio de S. Hajat fue creado con el propósito de permitir a cualquier persona interesada investigar hipótesis utilizando los datos ambientales y de salud disponibles, así como realizar análisis estadísticos y otros tipos de análisis adecuados, incluyendo la visualización de datos, para ello se utilizaron tecnologías de acceso abierto para su desarrollo [32]. Por último, Con el fin de satisfacer las demandas de calidad y revisión de datos, A. Pablo, A. N. Hoofnagle y P. C. Mathias han desarrollado un conjunto de herramientas utilizando software de CA. Este conjunto incluye una base de datos y un componente de visualización que permiten recopilar datos de manera efectiva [33].

TABLA V
ESTUDIOS SOBRE OBSERVATORIOS WEB

Descripción	Documentos
Observatorios que emplean CA	Yang et al [28] K. Haug et al [29] A. Neinstein et al[30] A. M. Fisher et al[31] S. Haja et al[32] A. Pablo et al[33]

c. ¿Cómo se mide la efectividad y calidad de los webs sites?

La medición de la efectividad y calidad de los observatorios web puede ser un proceso complejo, existen algunas métricas y enfoques comunes que usualmente se utilizan para evaluar la efectividad y calidad de las aplicaciones web. En los estudios revisados se puede observar que la ingeniería de requerimientos es un punto importante para identificar el alcance, los requerimientos funcionales y no funcionales que tendrá que cumplir un observatorio web. El análisis de requerimientos, conocido también como ingeniería de requerimientos, se refiere al procedimiento mediante el cual se obtienen las expectativas

de los usuarios o clientes en relación con las funcionalidades y la interacción de un software determinado, estas expectativas se traducen en requisitos específicos que guiarán el desarrollo del software [34]. Los requerimientos funcionales como no funcionales se definen el inicio del desarrollo para que al finalizar se verifique si se ha cumplido con los requerimientos planteados inicialmente y poder conocer si se desarrolló un software efectivo y de calidad. De los estudios revisados se puede observar requerimientos en común que comparten los observatorios web, A. Neinstein define los requerimientos que debe tener la base de datos del observatorio para almacenar y mostrar un amplio conjunto de información asociada para estudios en MetaboLights, arquitectura técnica para el desarrollo frontend y backend en el cual el frontend actúa como contenedor, mientras que el backend ejecuta el marco de modelado y proporciona los contenidos necesarios [29]. Estos dos componentes se combinan para ofrecer al usuario una aplicación funcional fácilmente accesible a través de la tecnología de Internet [35], lo que permite realizar búsqueda de temas para filtrar investigaciones por palabras claves, acceso y descarga de investigaciones, políticas de acceso y privacidad. Por su parte, A. Neinstein define requerimientos para el desarrollo como: construcción por piezas pequeñas y modulares, diseño web moderno, computación en la nube, desarrollo ágil, seguridad y privacidad [30]. Estos requerimientos se definen de acuerdo con las necesidades de cada observatorio web, pero existe un requerimiento primordial debido a que últimamente los dispositivos móviles se han convertido en la plataforma informática personal dominante [36]. Por ello, se debe garantizar la accesibilidad sin importar la plataforma, dispositivo o sistema operativo utilizado, lo que permite acceder al laboratorio no solo desde computadoras de escritorio y portátiles, sino también desde tabletas y teléfonos inteligentes [37]. Por otra parte, se mide poniendo a prueba al observatorio web o empleando herramientas que verifique que este sea efectivo y de calidad. Las pruebas de rendimiento se llevan a cabo empleando herramientas especializadas, como generadores de carga o herramientas de monitoreo. Durante estas pruebas, no solo se registran y analizan los tiempos de respuesta percibidos por el usuario, sino también se evalúa el uso de recursos desde la perspectiva del sistema [38]. La confidencialidad y la integridad de los datos no pueden ser aseguradas, por lo tanto, es necesario implementar mecanismos de autenticación y acceso más sólidos para garantizar la seguridad [39]. La seguridad en los sistemas es importante debido a que la seguridad de las aplicaciones web son generalmente débiles [40], emplear los escáneres de vulnerabilidades web son herramientas que automatizan el proceso de evaluar la seguridad de aplicaciones basadas en la web, permitiendo realizar análisis de seguridad a gran escala en múltiples aplicaciones web distintas [41], [42]. Por último, el estudio cuantitativo de I. M. Tarun muestra que se puede medir un software siendo utilizado por los usuarios en base a las puntuaciones Apdex [43].

TABLA VI
ESTUDIOS SOBRE LA EFECTIVIDAD Y CALIDAD DE UN
OBSERVATORIO WEB

DESCRIPCIÓN	DOCUMENTOS
La ingeniería de requerimientos en los observatorios web	A. Calabria et al [34] S. Ahmad and f. Hossain [35] A. Neinstein [30] R. Xu et al [39] S. Preehi [40]
Observatorios en diferentes dispositivos tecnológicos	L. Yuan et al [36] A. Villar-Martínez et al [37]
Herramientas para medir un observatorio web	V. Janani and k. Krishnamoorthy [38] M. Alsaleh et al [41] V. Garousi et al [42] I. M. Tarun [43]

IV. DISCUSIÓN

La revisión sistemática de literatura reveló estudios que abordaron el tema del uso de CA en observatorios web o webs sites. Estos estudios proporcionaron información valiosa sobre la efectividad de esta práctica. Al hacer uso del CA, el tema económico no se ve afectado debido a que para implementar soluciones de CA no se necesita de un capital inicial [18]. Asimismo, estas soluciones reducen el tiempo de elaboración de un proyecto ya que son ágiles, económicas y adaptables por lo que acelera el proceso [19], brindando la posibilidad de reutilizar componentes del sistema sin tener repercusiones [20]. Además, su implementación no restringe su rendimiento en general al crear una plataforma [23]. Al ser soluciones de CA su acceso es libre por lo que no se incurren en derechos de autor [24], y la mayoría de estas soluciones se pueden encontrar en plataformas como GitHub. También se encontró que ofrece escalabilidad y flexibilidad lo que permite adaptarse a medida que el proyecto evoluciona siendo posible el incremento de sus funcionalidades [25]. Con relación a los observatorios que emplearon soluciones de CA, se aprecian resultados exitosos, en el caso de MetaboLights, que permite compartir información relacionada sobre la metabolómica, como se mencionaba anteriormente sobre la ventaja económica que brindan las soluciones de CA, la necesidad de pagar por licencias y otros gastos no es necesario [17], en MetaboLights utilizaron herramientas como Apache Tomcat, MySQL y PostgreSQL [29], las cuales son de acceso abierto y no tienen costo alguno para ser utilizadas. Asimismo, la plataforma Tidepool para investigadores sobre diabetes, utilizó herramientas de CA para su implementación, lo cual hacer uso de ella brinda eficiencia y reducción de costos [30]. Se constató que existen métricas para evaluar la efectividad y calidad de un observatorio web, por una parte, se encuentra el análisis de los requerimientos, se refiere al proceso de identificar y comprender las expectativas de los usuarios o clientes, los cuales son previamente establecidos al inicio y que se tienen que cumplir para una correcta implementación [34]. También se encontró que existen requerimientos

indispensables como son la accesibilidad, seguridad, escalabilidad, confidencialidad e integridad, cabe resaltar que los requerimientos varían dependiendo de las necesidades de cada proyecto [36]. A diferencia de lo reportado, sobre la ingeniería de requerimientos para evaluar la efectividad y calidad, también existen las herramientas de monitoreo o generadores de carga son herramientas especializadas para evaluar el rendimiento y el uso de recursos del sistema [38]. Estas pruebas permiten identificar y solucionar problemas de rendimiento antes de que afecten la experiencia del usuario y garantizan que el observatorio web pueda manejar una carga de trabajo realista. Además, registran métricas como el tiempo de respuesta, el uso de recursos y otros indicadores clave, lo que permite monitorear el rendimiento del observatorio web de manera continua y detectar cualquier anomalía o problema. En cuanto al estudio de I.M Tarun destaca la utilidad de las puntuaciones Apdex como una medida para evaluar el rendimiento de un software en base a su uso por parte de los usuarios [43]. Las puntuaciones Apdex proporcionan una forma cuantitativa de medir la calidad y efectividad de un software en función de la experiencia del usuario. Sin embargo, es importante destacar que las puntuaciones Apdex son una medida cuantitativa que puede tener limitaciones, ya que no evalúan todos los aspectos del software, como la usabilidad, seguridad, escalabilidad entre otros requerimientos no funcionales.

V. CONCLUSIONES

La innovación tecnológica de la ingeniería de software tiene como uno de sus pilares el uso de software de código abierto. El 98% de las empresas que utilizan software de código abierto reducen el tiempo de elaboración de un proyecto, logrando agilidad en los procesos, economía en los recursos y adaptabilidad tecnológica; al reutilizar componentes del sistema, sin tener repercusiones negativas. En términos de factores influyentes, existen requerimientos indispensables como son la accesibilidad, seguridad, escalabilidad, confidencialidad e integridad; cabe resaltar que los requerimientos varían dependiendo de las necesidades de cada proyecto. En tal sentido, es importante el desarrollo de investigaciones que permitan determinar las diversas formas en que el código abierto puede ser utilizado para proyectos de software.

VI. REFERENCIAS

- [1] B. Wijnen, E. J. Hunt, G. C. Anzalone, and J. M. Pearce, "Open-source syringe pump library," *PLoS One*, vol. 9, no. 9, Sep. 2014.
- [2] G. M. Parra Mora and H. P. Serna González, "La descripción documental como medio de acceso a la memoria institucional: experiencia de uso del software ICA AtoM en la Fundación Universitaria para el Desarrollo Humano –UNINPAHU– Bogotá, Colombia," *e-Ciencias de la Información*, Jan. 2020.
- [3] "Dialnet-TraducciónDeConocimientosDelSoftwareLibreYDeCodigo-7778053 (3)".
- [4] "Open-Source-Platform-and-Elearning-in-a-Public-University-in-PeruJournal-of-Pharmaceutical-Negative-Results".

- [5] E. DE Nacional Biblioteconomía Y Archivonomía and O. Arriola Navarrete Mtra Emma Hernández Gómez MÉXICO, "SECRETARÍA DE EDUCACIÓN PÚBLICA DIRECCIÓN GENERAL DE EDUCACIÓN SUPERIOR UNIVERSITARIA," 2011.
- [6] T. Leinonen, Ç. Çöltekin, and J. Nerbonne, "Using Gabmap," *Lingua*, vol. 178, pp. 71–83, Jul. 2016.
- [7] Y. Wu, S. C. H. Hoi, C. Liu, J. Lu, D. Sahoo, and N. Yu, "SOL: A library for scalable online learning algorithms," *Neurocomputing*, vol. 260, pp. 9–12, Oct. 2017.
- [8] D. Han, J. Pan, X. Zhao, and W. Chen, "NetV.js: A web-based library for high-efficiency visualization of large-scale graphs and networks," *Visual Informatics*, vol. 5, no. 1, pp. 61–66, Mar. 2021.
- [9] S. Hajat et al., "Development of a browser application to foster research on linking climate and health datasets: Challenges and opportunities," *Science of the Total Environment*, vol. 575, pp. 79–86, Jan. 2017.
- [10] Z. Chen et al., "Enhancing example-based code search with functional semantics," *Journal of Systems and Software*, vol. 165, Jul. 2020.
- [11] L. Yuan, J. Ren, L. Gao, Z. Tang, and Z. Wang, "Using Machine Learning to Optimize Web Interactions on Heterogeneous Mobile Systems," *IEEE Access*, vol. 7, pp. 139394–139408, 2019.
- [12] W. Mao, B. Sun, G. Xu, C. Liu, C. Si, and W. Wang, "Understanding Effects of Collaborations in Developing Mobile Computing Systems: Popularity, Efficiency, and Quality," *IEEE Access*, vol. 7, pp. 33380–33392, 2019.
- [13] D. Carrizo and C. Moller, "Estructuras metodológicas de revisiones sistemáticas de literatura en Ingeniería de Software: un estudio de mapeo sistemático Methodological structures of systematic literature review in software engineering: a systematic mapping study."
- [14] M. J. Page et al., "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," *Syst Rev*, vol. 10, no. 1, Dec. 2021.
- [15] S. Zhang and Y. Li, "Modeling and Simulation Study of Designer's Bidirectional Behavior of Task Selection in Open Source Design Process," *Math Probl Eng*, vol. 2017, 2017.
- [16] K. S. Burgdorf et al., "Digital questionnaire platform in the Danish Blood Donor Study," *Comput Methods Programs Biomed*, vol. 135, pp. 101–104, Oct. 2016.
- [17] D. Zhang, X. Chen, and H. Yao, "Development of a prototype web-based decision support system for watershed management," *Water (Switzerland)*, vol. 7, no. 2, pp. 780–793, 2015.
- [18] J. Pérez-Padillo, J. G. Morillo, J. Ramirez-Faz, M. T. Roldán, and P. Montesinos, "Design and implementation of a pressure monitoring system based on iot for water supply networks," *Sensors (Switzerland)*, vol. 20, no. 15, pp. 1–19, Aug. 2020.
- [19] H. Kim, W. K. Jung, I. G. Choi, and S. H. Ahn, "A low-cost vision-based monitoring of computer numerical control (CNC) machine tools for small and medium-sized enterprises (SMES)," *Sensors (Switzerland)*, vol. 19, no. 20, Oct. 2019.
- [20] C. Yuan et al., "Criteria2Query: A natural language interface to clinical databases for cohort definition," *Journal of the American Medical Informatics Association*, vol. 26, no. 4, pp. 294–305, Feb. 2019.
- [21] S. Wang et al., "Scalable and accessible personalized photodynamic therapy optimization with FullMonte and PDT-SPACE," *J Biomed Opt*, vol. 27, no. 08, Apr. 2022.
- [22] Y. V. Zaytsev and A. Morrison, "Increasing quality and managing complexity in neuroinformatics software development with continuous integration," *Front Neuroinform*, vol. 6, no. DEC, Jan. 2013.
- [23] D. Amaxilatis, O. Akrivopoulos, G. Mylonas, and I. Chatzigiannakis, "An IoT-based solution for monitoring a fleet of educational buildings focusing on energy efficiency," *Sensors (Switzerland)*, vol. 17, no. 10, Oct. 2017.
- [24] D. A. Ettlin et al., "Design, construction, and technical implementation of a web-based interdisciplinary symptom evaluation (WISE) - a heuristic proposal for orofacial pain and temporomandibular disorders," *Journal of Headache and Pain*, vol. 17, no. 1, Dec. 2016.
- [25] C. T. Yang, S. T. Chen, Y. W. Chan, and Y. C. Shen, "On construction of a cloud storage system with heterogeneous software-defined storage technologies," *Human-centric Computing and Information Sciences*, vol. 9, no. 1, Dec. 2019.
- [26] Z. Chen et al., "Enhancing example-based code search with functional semantics," *Journal of Systems and Software*, vol. 165, Jul. 2020.
- [27] A. Fabregat et al., "Reactome graph database: Efficient access to complex pathway data," *PLoS Comput Biol*, vol. 14, no. 1, Jan. 2018.
- [28] P. Yang, H. Fang, and J. Lin, "Anserini: Reproducible ranking baselines using lucene," *Journal of Data and Information Quality*, vol. 10, no. 4, Oct. 2018.
- [29] K. Haug et al., "MetaboLights - An open-access general-purpose repository for metabolomics studies and associated meta-data," *Nucleic Acids Res*, vol. 41, no. D1, Jan. 2013.
- [30] A. Neinstein et al., "A case study in open source innovation: Developing the Tidepool Platform for interoperability in type 1 diabetes management," *Journal of the American Medical Informatics Association*, vol. 23, no. 2, pp. 324–332, Mar. 2016.
- [31] A. M. Fisher et al., "User-centered design and usability testing of RxMAGIC: a prescription management and general inventory control system for free clinic dispensaries," *BMC Health Serv Res*, vol. 18, no. 1, p. 703, Sep. 2018.
- [32] S. Hajat et al., "Development of a browser application to foster research on linking climate and health datasets: Challenges and opportunities," *Science of the Total Environment*, vol. 575, pp. 79–86, Jan. 2017.
- [33] A. Pablo, A. N. Hoofnagle, and P. C. Mathias, "Listening to your mass spectrometer: An open-source toolkit to visualize mass spectrometer data," *Journal of Mass Spectrometry and Advances in the Clinical Lab*, vol. 23, pp. 44–49, Jan. 2022.
- [34] A. Calabria, G. Spinozzi, F. Benedicenti, E. Tenderini, and E. Montini, "adLIMS: A customized open source software that allows bridging clinical and basic molecular research studies," *BMC Bioinformatics*, vol. 16, Jun. 2015.
- [35] S. K. Ahmad and F. Hossain, "A web-based decision support system for smart dam operations using weather forecasts," *Journal of Hydroinformatics*, vol. 21, no. 5, pp. 687–707, Sep. 2019.
- [36] L. Yuan, J. Ren, L. Gao, Z. Tang, and Z. Wang, "Using Machine Learning to Optimize Web Interactions on Heterogeneous Mobile Systems," *IEEE Access*, vol. 7, pp. 139394–139408, 2019.
- [37] A. Villar-Martinez, L. Rodriguez-Gil, I. Angulo, P. Orduna, J. Garcia-Zubia, and D. Lopez-De-Ipina, "Improving the Scalability and Replicability of Embedded Systems Remote Laboratories through a Cost-Effective Architecture," *IEEE Access*, vol. 7, pp. 164164–164185, 2019.
- [38] V. Janani and K. Krishnamoorthy, "Evaluation of cloud based performance testing for online shopping websites," *Indian J Sci Technol*, vol. 8, no. 35, Dec. 2015.
- [39] R. Xu, W. Jin, and D. Kim, "Microservice security agent based on API gateway in edge computing," *Sensors (Switzerland)*, vol. 19, no. 22, Nov. 2019.
- [40] S. Preethi, "Analysis of Vulnerability Detection Tool for Web Services," 2018. [Online]. Available: <http://www.tpc.org>
- [41] M. Alsaleh, N. Alomar, M. Alshreef, A. Alarifi, and A. M. Al-Salman, "Performance-Based Comparative Assessment of Open Source Web Vulnerability Scanners," *Security and Communication Networks*, vol. 2017, 2017.
- [42] V. Garousi, A. B. Keleş, Y. Balaman, Z. Ö. Güler, and A. Arcuri, "Model-based testing in practice: An experience report from the web applications domain," *Journal of Systems and Software*, vol. 180, Oct. 2021.
- [43] I. M. Tarun, "The effectiveness of a customized online collaboration tool for teaching and learning," *Journal of Information Technology Education: Research*, vol. 18, pp. 275–292, 2019.