



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

CARRERA DE INGENIERÍA EN SISTEMAS E INFORMÁTICA

**TRABAJO DE TITULACIÓN PREVIO LA OBTENCIÓN DEL TÍTULO
DE INGENIERO EN SISTEMAS E INFORMÁTICA**

**TEMA: ESTUDIO DE CASO PARA LA ADOPCIÓN DEL PARADIGMA
DE INVESTIGACIÓN REPRODUCIBLE EN INGENIERÍA DE
SOFTWARE**

**AUTORES: FERNÁNDEZ PEÑAFIEL, SANTIAGO MIGUEL,
RODRÍGUEZ QUINTO, MARCO VINICIO**

DIRECTOR: ING. FONSECA C, EFRAÍN RODRIGO

SANGOLQUÍ

2019

CERTIFICADO



DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN
CARRERA DE INGENIERÍA EN SISTEMAS E INFORMÁTICA

CERTIFICACIÓN

Certifico que el trabajo de titulación, “*ESTUDIO DE CASO PARA LA ADOPCIÓN DEL PARADIGMA DE INVESTIGACIÓN REPRODUCIBLE EN INGENIERÍA DE SOFTWARE*” fue realizado por los señores *FERNÁNDEZ PEÑAFIEL SANTIAGO MIGUEL* y *RODRÍGUEZ QUINTO MARCO VINICIO*, el mismo ha sido revisado en su totalidad, analizado por la herramienta de verificación de similitud de contenido; por lo tanto cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de Fuerzas Armadas ESPE, razón por la cual me permito acreditar y autorizar para que lo sustenten públicamente.

Sangolquí, 16 de enero del 2019

Dr. Efraín R. Fonseca C.

DIRECTOR

AUTORÍA DE RESPONSABILIDAD



DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

CARRERA DE INGENIERÍA EN SISTEMAS E INFORMÁTICA

AUTORÍA DE RESPONSABILIDAD

Nosotros, **FERNÁNDEZ PEÑAFIEL SANTIAGO MIGUEL** y **RODRÍGUEZ QUINTO MARCO VINICIO**, declaramos que el contenido, ideas y criterios del trabajo de titulación: **“ESTUDIO DE CASO PARA LA ADOPCIÓN DEL PARADIGMA DE INVESTIGACIÓN REPRODUCIBLE EN INGENIERÍA DE SOFTWARE”** es de nuestra autoría y responsabilidad, cumpliendo con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de Fuerzas Armadas ESPE, respetando los derechos intelectuales de terceros y referenciando las citas bibliográficas.

Consecuentemente el contenido de la investigación mencionada es veraz.

Sangolquí, 16 de enero del 2019

FERNÁNDEZ PEÑAFIEL
SANTIAGO MIGUEL
C.C. 1725211302

RODRÍGUEZ QUINTO
MARCO VINICIO
C.C. 1720215308

AUTORIZACIÓN



DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN
CARRERA DE INGENIERÍA EN SISTEMAS E INFORMÁTICA

AUTORIZACIÓN

Nosotros, **FERNÁNDEZ PEÑAFIEL SANTIAGO MIGUEL** y **RODRÍGUEZ QUINTO MARCO VINICIO**, autorizamos a la Universidad de las Fuerzas Armadas ESPE publicar el trabajo de titulación: **“ESTUDIO DE CASO PARA LA ADOPCIÓN DEL PARADIGMA DE INVESTIGACIÓN REPRODUCIBLE EN INGENIERÍA DE SOFTWARE”** en el Repositorio Institucional, cuyo contenido, ideas y criterios son de nuestra autoría y responsabilidad.

Sangolquí, 16 de enero del 2019

FERNÁNDEZ PEÑAFIEL
SANTIAGO MIGUEL
C.C. 1725211302

RODRÍGUEZ QUINTO
MARCO VINICIO
C.C. 1720215308

DEDICATORIA

Esta tesis se la dedicamos a Dios, que con su guía celestial nos ha permitido cumplir nuestros objetivos y nos ha enseñado a encarar las adversidades sin desfallecer en el intento.

A nuestros padres que son los pilares fundamentales de nuestras vidas, que con su apoyo y consejos nos han motivado para cumplir con nuestros ideales; y quienes han estado presentes en los momentos difíciles y de alegría.

A nuestros amigos quienes con su apoyo y palabras de aliento nos han impulsado a culminar la formación universitaria. A nuestro compañero de tesis que con su permanente apoyo y comprensión ha permitido alcanzar nuestro objetivo tan deseado.

AGRADECIMIENTO

Queremos agradecer primero a Dios, por habernos brindado la sabiduría y fortaleza necesaria para culminar con éxito nuestra carrera universitaria, y nos ha permitido superar y afrontar las adversidades que se nos han presentado.

A nuestros padres y hermanos que son nuestros eternos consejeros, quienes con su gran amor, apoyo incondicional y paciencia nos guiaron por el camino del bien y nos enseñaron a luchar por nuestros objetivos.

Al Ingeniero Efraín Fonseca, quien con su guía y dedicación nos impartió el conocimiento y facilitó el tiempo necesario para culminar el proyecto de titulación.

A nuestros amigos y confidentes con los que hemos compartido grandiosos e inolvidables momentos y que han sido una gran compañía para lograr esta meta.

Finalmente, a todas aquellas personas que de una u otra forma han contribuido a que este sueño se haga realidad.

ÍNDICE DE CONTENIDOS

| | |
|--|-------------|
| CERTIFICADO | i |
| AUTORÍA DE RESPONSABILIDAD..... | ii |
| AUTORIZACIÓN | iii |
| DEDICATORIA | iv |
| AGRADECIMIENTO | v |
| ÍNDICE DE CONTENIDOS..... | vi |
| ÍNDICE DE TABLAS. | xi |
| ÍNDICE DE FIGURAS | xiii |
| RESUMEN..... | xv |
| ABSTRACT..... | xvi |
| ACRÓNIMOS..... | xvii |
| CAPÍTULO I..... | 1 |
| INTRODUCCIÓN | 1 |
| 1.1 ANTECEDENTES | 1 |
| 1.2 PROBLEMÁTICA | 2 |
| 1.1 JUSTIFICACIÓN..... | 3 |
| 1.2 OBJETIVOS..... | 4 |
| 1.2.1 Objetivo General | 4 |
| 1.2.2 Objetivos Específicos | 4 |
| 1.3 ALCANCE | 5 |
| 1.4 HIPÓTESIS | 6 |
| CAPÍTULO II | 7 |
| MARCO METODOLÓGICO | 7 |
| 2.1. MARCO TEÓRICO..... | 7 |
| 2.1.1. Señalamiento de variables..... | 7 |

| | |
|--|----|
| 2.1.1.1. Red de categorías..... | 7 |
| 2.1.2. Ciencia | 8 |
| 2.1.3. Investigación Científica | 9 |
| 2.1.3.1. Objetivos de la investigación. | 9 |
| 2.1.3.2. Características de la Investigación Científica..... | 10 |
| 2.1.3.3. Tipos de Investigación Científica..... | 10 |
| 2.1.4. Método Científico..... | 13 |
| 2.1.5. Investigación Reproducible | 13 |
| 2.1.5.1. Etapas básicas en Investigación Reproducible..... | 15 |
| 2.1.6. Ciencias De La Computación | 16 |
| 2.1.6.1. Algoritmos y estructuras de datos | 17 |
| 2.1.6.2. Lenguajes de programación..... | 17 |
| 2.1.6.3. Arquitectura | 17 |
| 2.1.6.4. Computación numérica y simbólica | 17 |
| 2.1.6.5. Sistemas Operativas..... | 18 |
| 2.1.6.6. Metodología e ingeniería de software | 18 |
| 2.1.6.7. Bases de datos y recuperación de información | 18 |
| 2.1.6.8. Inteligencia artificial y robótica | 18 |
| 2.1.6.9. Comunicación humano computadora | 18 |
| 2.1.7. Ingeniería de Software..... | 18 |
| 2.1.8. Ingeniería de Software Empírica | 19 |
| 2.1.8.1. Experimento controlado | 19 |
| 2.1.8.2. Caso de estudio..... | 20 |
| 2.1.8.3. Encuesta | 20 |
| 2.1.8.4. Etnografía | 20 |
| 2.1.8.5. Investigación de acción | 21 |

| | | |
|---|---|-----------|
| 2.2. | ESTADO DEL ARTE | 21 |
| 2.2.1. | Planteamiento de la revisión de literatura | 21 |
| 2.2.2. | Conformación del grupo de control (GC) | 22 |
| 2.2.3. | Construcción y afinación de la cadena de búsqueda | 22 |
| 2.2.4. | Selección de estudios | 23 |
| 2.2.5. | Determinación de las características del estado del arte..... | 24 |
| 2.2.5.1. | Resumen de los estudios primarios..... | 24 |
| 2.2.5.2. | Características del estado del arte | 28 |
| 2.3. | METODOLOGÍA DE LA INVESTIGACIÓN | 28 |
| 2.3.1. | Investigación Exploratoria De Método Inductivo | 28 |
| 2.3.2. | Metodología Design Science | 29 |
| 2.3.3. | Ciclo de diseño e ingeniería en Design Science | 30 |
| 2.3.3.1. | Investigación del problema | 30 |
| 2.3.3.2. | Diseño del tratamiento..... | 31 |
| 2.3.3.3. | Validación del tratamiento..... | 32 |
| 2.3.3.4. | Implementación del tratamiento | 32 |
| 2.3.3.5. | Evaluación de la implementación | 33 |
| 2.3.4. | Design Science aplicado | 34 |
| CAPÍTULO III..... | | 36 |
| MODELOS PROPUESTOS PARA REPRODUCIBILIDAD | | 36 |
| 3.1. | MODELO CONCEPTUAL DE INVESTIGACIÓN REPRODUCIBLE | 36 |
| 3.2. | MODELO DE PROCESOS PARA INVESTIGACIÓN REPRODUCIBLE..... | 38 |
| 3.2.1. | Pools | 40 |
| 3.2.2. | Lanes | 40 |
| 3.2.3. | Subprocesos | 41 |
| 3.2.3.1. | Diseñar Estudio | 41 |

| | |
|--|-----------|
| 3.2.3.2. Replicar Estudio | 44 |
| 3.2.3.3. Analizar Conjunto Global de Datos..... | 52 |
| CAPÍTULO IV | 54 |
| DISEÑO DEL PROTOTIPO DE PLATAFORMA WEB | 54 |
| 4.1 DISEÑO DE BASE DE DATOS | 54 |
| 4.1.1. Modelo conceptual | 54 |
| 4.1.2. Modelo lógico | 55 |
| 4.1.3. Modelo físico..... | 55 |
| 4.2 SCRUM | 56 |
| 4.2.1 Historias de Usuario | 57 |
| 4.3 REQUISITOS DEL PROTOTIPO | 62 |
| 4.3.1 Especificación de requisitos | 62 |
| 4.3.1.1. Requisitos funcionales | 62 |
| 4.3.1.2. Requisitos no funcionales..... | 70 |
| 4.3.2 Casos de uso | 71 |
| 4.4 ARQUITECTURA DE SOFTWARE | 73 |
| 4.4.1. Arquitectura plataforma web | 73 |
| 4.4.2. Arquitectura programa standalone | 76 |
| 4.5 INTERFAZ DE USUARIO..... | 76 |
| CAPÍTULO V | 83 |
| APLICACIÓN DEL PROTOTIPO AL CONTEXTO/PROBLEMA | 83 |
| 5.1 CASO DE ESTUDIO | 83 |
| 5.1.1. Proyecto NaPiRE..... | 83 |
| 5.1.2. Descripción de la encuesta NaPiRE | 83 |
| 5.2 VALIDACIÓN DE LA PROPUESTA | 84 |
| 5.2.1. Diseño de la encuesta | 85 |

| | |
|---|------------|
| 5.2.1.1. Encuesta de satisfacción de usuario..... | 87 |
| 5.2.2. Proceso de ejecución de encuestas | 91 |
| 5.2.2.1. Instrucciones de uso de la herramienta | 91 |
| 5.2.3. Análisis de resultados | 97 |
| 5.2.3.1. Universidad a la que pertenece | 97 |
| 5.2.3.2. Pregunta 1 | 98 |
| 5.2.3.3. Pregunta 2..... | 99 |
| 5.2.3.4. Pregunta 3..... | 100 |
| 5.2.3.5. Pregunta 4..... | 102 |
| CAPÍTULO VI | 106 |
| CONCLUSIONES, RECOMENDACIONES Y TRABAJOS FUTUROS | 106 |
| 6.1 CONCLUSIONES | 106 |
| 6.2 RECOMENDACIONES | 107 |
| 6.3 TRABAJOS FUTUROS..... | 107 |

ÍNDICE DE TABLAS

| | |
|---|----|
| Tabla 1 Preguntas de Investigación | 5 |
| Tabla 2 Grupo de Control | 22 |
| Tabla 3 Estudio Seleccionados | 24 |
| Tabla 4 Historia de usuario 1 | 57 |
| Tabla 5 Historia de usuario 2 | 57 |
| Tabla 6 Historia de usuario 3 | 57 |
| Tabla 7 Historia de usuario 4 | 58 |
| Tabla 8 Historia de usuario 5 | 58 |
| Tabla 9 Historia de usuario 6 | 59 |
| Tabla 10 Historia de usuario 7 | 59 |
| Tabla 11 Historia de usuario 8 | 59 |
| Tabla 12 Historia de usuario 9 | 60 |
| Tabla 13 Historia de usuario 10 | 60 |
| Tabla 14 Historia de usuario 11 | 61 |
| Tabla 15 Historia de usuario 12 | 61 |
| Tabla 16 Historia de usuario 13 | 61 |
| Tabla 17 Requisito funcional 1 | 63 |
| Tabla 18 Requisito funcional 2 | 64 |
| Tabla 19 Requisito funcional 3 | 64 |
| Tabla 20 Requisito funcional 4 | 65 |
| Tabla 21 Requisito funcional 5 | 66 |
| Tabla 22 Requisito funcional 6 | 66 |
| Tabla 23 Requisito funcional 7 | 67 |

| | |
|---|-----|
| Tabla 24 <i>Requisito funcional 8</i> | 68 |
| Tabla 25 <i>Requisito funcional 9</i> | 68 |
| Tabla 26 <i>Requisito funcional 10</i> | 69 |
| Tabla 27 <i>Requisito funcional 11</i> | 69 |
| Tabla 28 <i>Requisito no funcional 1</i> | 70 |
| Tabla 29 <i>Requisito no funcional 2</i> | 70 |
| Tabla 30 <i>Requisito no funcional 3</i> | 71 |
| Tabla 31 <i>Requisito no funcional 4</i> | 71 |
| Tabla 32 <i>Correspondencia de preguntas</i> | 86 |
| Tabla 33 <i>Resultados universidad a la que pertenecen los encuestados</i> | 97 |
| Tabla 34 <i>Resultados primera pregunta encuesta</i> | 98 |
| Tabla 35 <i>Resultados segunda pregunta encuesta</i> | 99 |
| Tabla 36 <i>Resultados tercera pregunta encuesta</i> | 101 |
| Tabla 37 <i>Resultados cuarta pregunta encuesta literal a</i> | 102 |
| Tabla 38 <i>Resultados cuarta pregunta encuesta literal b</i> | 103 |
| Tabla 39 <i>Resultados cuarta pregunta encuesta literal c</i> | 103 |
| Tabla 40 <i>Resultados cuarta pregunta encuesta literal d</i> | 104 |
| Tabla 41 <i>Resultados cuarta pregunta encuesta literal e</i> | 104 |

ÍNDICE DE FIGURAS

| | |
|--|----|
| <i>Figura 1.</i> Red de categorías de las variables de investigación | 7 |
| <i>Figura 2.</i> Método propuesto para la elaboración del estado del arte..... | 21 |
| <i>Figura 3.</i> El ciclo de ingeniería | 30 |
| <i>Figura 4.</i> Ciclo Design Science aplicado a la investigación | 35 |
| <i>Figura 5.</i> Modelo conceptual de investigación reproducible | 38 |
| <i>Figura 6.</i> Proceso de Investigación Reproducible..... | 40 |
| <i>Figura 7.</i> Subproceso Diseñar Estudio | 43 |
| <i>Figura 8.</i> Subproceso Replicar Estudio..... | 44 |
| <i>Figura 9.</i> Subproceso Solicitar Autorización | 46 |
| <i>Figura 10.</i> Subproceso Crear Cuenta de Usuario | 47 |
| <i>Figura 11.</i> Subproceso Solicitar Replicación..... | 48 |
| <i>Figura 12.</i> Subproceso Ejecutar Replicación de Estudio | 51 |
| <i>Figura 13.</i> Subproceso Analizar Conjunto Global de Datos | 53 |
| <i>Figura 14.</i> Modelo conceptual de base de datos | 54 |
| <i>Figura 15.</i> Modelo lógico de base de datos..... | 55 |
| <i>Figura 16.</i> Modelo físico de base de datos..... | 56 |
| <i>Figura 17.</i> Diagrama de casos de uso..... | 72 |
| <i>Figura 18.</i> Arquitectura de software..... | 73 |
| <i>Figura 19.</i> Arquitectura plataforma web | 75 |
| <i>Figura 20.</i> Arquitectura programa standalone..... | 76 |
| <i>Figura 21.</i> Interfaz de usuario Login..... | 77 |
| <i>Figura 22.</i> Interfaz de usuario gestión de solicitudes..... | 78 |
| <i>Figura 23.</i> Interfaz de usuario gestión de estudios..... | 79 |

| | |
|---|-----|
| Figura 24. Interfaz de usuario diseño de estudio..... | 79 |
| Figura 25. Interfaz de usuario replicaciones | 80 |
| Figura 26. Interfaz de usuario categorizar respuestas | 81 |
| Figura 27. Interfaz de usuario resultados | 82 |
| Figura 28. Interfaz de usuario encuesta..... | 82 |
| Figura 29. Actividad de login..... | 92 |
| Figura 30. Actividad elegir estudio para replicar..... | 93 |
| Figura 31. Actividad de difusión..... | 94 |
| Figura 32. Actividad de análisis de resultados..... | 95 |
| Figura 33. Actividad respaldar resultados desde plataforma | 96 |
| Figura 34. Actividad respaldar resultados desde standalone | 96 |
| Figura 35. Universidad a la que pertenecen los encuestados | 98 |
| Figura 36. Razones por las que la herramienta apoya al caso de estudio | 99 |
| Figura 37. Fases de la investigación agilizados por la herramienta..... | 100 |
| Figura 38. Formas en que ayuda la herramienta en el caso de estudio | 101 |
| Figura 39. Beneficios de la herramienta según escala de Likert..... | 105 |

RESUMEN

La Ingeniería de Software ha tenido un aumento significativo en el número de estudios empíricos en las últimas décadas. Mientras los estudios empíricos se hacen más comunes en la Ingeniería de Software, la importancia de que estos sean reproducibles aumenta. El problema surge al intentar sintetizar esta gran cantidad de trabajos ya que puede ser una tarea ardua, principalmente por la dificultad que implica la verificación de los resultados de las investigaciones existentes, por otra parte, replicar estudios puede llegar a ser muy agotador si no se los ha diseñado con este propósito debido a que no se publican detalles profundos utilizados en la investigación. Los beneficios de utilizar el paradigma de investigación reproducible repercuten en que las repeticiones ayudan a aumentar el cuerpo de conocimiento alrededor de Ingeniería de Software, lo que a su vez conduce a una mayor madurez del campo. Una replicación de un estudio también viene con beneficios para el estudio original en términos de aumento de la confianza para el estudio conducido y los resultados informados. Es por esto que con el motivo de integrar el paradigma de investigación reproducible a la ingeniería de software se desarrolló el prototipo de investigación reproducible aplicado al caso de estudio NaPiRE, el cual busca facilitar la replicación de investigaciones de tipo encuesta. Como resultado se propusieron modelos que incorporen el paradigma de investigación reproducible en la ingeniería de software aplicado a investigaciones de tipo encuesta, mismos que resultaron factibles para desarrollar una herramienta que facilite el proceso de replicación y contribuya a validar y corroborar los hallazgos conseguidos con el caso de estudio.

PALABRAS CLAVE:

- **INVESTIGACIÓN REPRODUCIBLE**
- **REPLICACIÓN**
- **REPRODUCIBILIDAD**
- **HERRAMIENTA**
- **PROTOTIPO**

ABSTRACT

Software Engineering has had a significant increase in the number of empirical studies in recent decades. While empirical studies become more common in Software Engineering, the importance of reproducibility increases. The problem arises when trying to synthesize this great amount of works since it can be an arduous task, mainly for the difficulty that implies the verification of the results of the existing investigations, on the other hand, to replicate studies can get to be very exhausting if it is not he has designed them for this purpose because they do not publish deep details used in the research. The benefits of using the reproducible research paradigm have repercussions in that the repetitions help to increase the body of knowledge around Software Engineering, which in turn leads to a greater maturity of the field. A replication of a study also comes with benefits for the original study in terms of increased confidence for the study conducted and the results reported. That is why, in order to integrate the paradigm of reproducible research into software engineering, the prototype of reproducible research applied to the NaPiRE study case was developed, which seeks to facilitate the replication of survey-type investigations. As a result, models were proposed that incorporate the paradigm of reproducible research in software engineering applied to survey-type investigations, which were feasible to develop a tool that facilitates the replication process and contributes to validate and corroborate the findings obtained with the case study.

KEYWORDS:

- **REPRODUCIBLE RESEARCH**
- **REPLICATION**
- **REPRODUCIBILITY**
- **TOOL**
- **PROTOTYPE**

ACRÓNIMOS

| | |
|--------------------|--|
| GrIMPSoftII | Grupo de Investigación en Modelos de Producción de Software. |
| NaPiRE | Naming the Pain in Requirements Engineering, red de investigación conformado por más de 25 de países la cual tiene como objetivo establecer una teoría general sobre las prácticas y problemas de la ingeniería de requisitos. |

CAPÍTULO I

INTRODUCCIÓN

1.1 ANTECEDENTES

La Ingeniería de Software ha tenido un aumento significativo en el número de estudios empíricos en las últimas décadas, lo que indica que actualmente es un campo de interés en continuo crecimiento (Larsson, Lindqvist, & Torkar, 2014). Según (Zendler, 2001), este incremento inicia en la década de los noventa a raíz del apareamiento de diversos laboratorios financiados para investigar en el campo de la ingeniería de software.

El paradigma de investigación reproducible tuvo sus inicios en los años noventa, de la mano de Jon Claerbout, un investigador de la Universidad de Stanford, quien junto a su equipo desarrollaron un sistema basado en scripts para obtener las figuras de un documento junto con la información enlazada a cada figura y poder reproducirlas en conjunto en otros documentos, usando un comando en Unix; es decir, cualquier persona podía reproducir resultados previos, para utilizarlos en su trabajo (Claerbout, 1991). Claerbout identificó tres tipos de resultados: aquellos fácilmente reproducibles, resultados de reproducción condicionada y los resultados que no se pueden reproducir. Posteriormente el sistema basado en Makefile de Claerbout y su equipo fue implementado y mejorado en el proyecto Madagascar utilizando SCons, que es un lenguaje similar a Makefiles pero que permite reproducir de manera más sencilla y es multiplataforma (Fomel, Sava, Vlad, Liu, & Bashkardin, 2013).

La investigación reproducible ha sido ampliamente usada en distintas disciplinas como, por ejemplo: medicina (Iqbal, Wallach, Khoury, & Ioannidis, 2016), farmacología (Iqbal et al., 2016) y biología (Peng, 2011). Sin embargo, pese a que la Ingeniería de Software es una ciencia dedicada a la recolección de información, formular nuevas teorías y compartir sus resultados con otros científicos, el paradigma de investigación reproducible no ha sido ampliamente considerado en esta disciplina. Como consecuencia existe dificultad al momento de evaluar la

credibilidad y confiabilidad de los resultados de investigaciones en el campo (Fomel, 2015), debido a que no existe un intercambio abierto de información que permita que nuevos investigadores recreen los estudios propuestos en la ingeniería de software.

A partir del año 1994, Da Silva manifiesta que se puede observar una relación entre la investigación reproducible y la ingeniería de software con la primera publicación que se trata explícitamente de la reproducción de un estudio empírico (da Silva et al., 2012), lo que demuestra que el tema se ha trabajado aproximadamente durante dos décadas. El interés de la comunidad investigativa hacia el tema se ha manifestado mediante la realización de un gran evento denominado “The International Workshop on Replication in Empirical Software Engineering Research”, que tuvo dos ediciones en el año 2010 y 2011. Sin embargo considerando la importancia de la replicación en las ciencias empíricas en general y en específico para la ingeniería de software no existe una base de conocimiento sólida acerca de la investigación reproducible que provea definiciones claras y no ambiguas a las principales interrogantes con respecto al tema como ¿Qué es exactamente reproducir un estudio?, ¿Qué es exactamente una reproducción exitosa?, entre otras planteadas por (Schmidt, 2009).

1.2 PROBLEMÁTICA

La cantidad de trabajos relacionados a la Ingeniería de software lleva aumentando desde los noventas (Zendler, 2001); sin embargo, como se puede observar en el trabajo de (Sjøberg et al., 2005) o (Fernández, 2011) sintetizar esta gran cantidad de trabajos puede ser una tarea ardua, principalmente por la dificultad que implica la verificación de los resultados de las investigaciones existentes.

El problema radica en la falta de métodos, procesos o modelos que permitan duplicar, validar e inclusive cuestionar los resultados de las investigaciones de Ingeniería de Software por parte de la comunidad científica, por otra parte la diversidad de tipos de estudios dificulta

el establecimiento de un modelo o método genérico que permita la reproducibilidad de investigaciones (Sjøberg et al., 2005). Teniendo en cuenta las características de los estudios candidatos a ser replicados se destaca que deben tener todos sus datos accesibles y de igual modo, todo proceso aplicado al estudio debe ser replicable (Larsson et al., 2014). Como consecuencia, muchas de estas investigaciones se queden aisladas sin aportar ningún tipo conocimiento científico a la Ingeniería de Software lo que dificulta que esta madure como ciencia por la falta de rigurosidad del conocimiento (Sjøberg et al., 2005).

1.1 JUSTIFICACIÓN

La Sociedad Estadounidense de Física (APS por sus siglas en inglés American Physical Society) indica que: “Science is the systematic enterprise of gathering knowledge about the universe and organizing and condensing that knowledge into testable laws and theories. The success and credibility of science are anchored in the willingness of scientists to expose their ideas and results to independent testing and replication by other scientists. This requires the complete and open exchange of data, procedures and materials.”(Cassidy, Holton, & Retherford, 2013). Dentro de cada disciplina, la investigación ha conducido a hallazgos trascendentales basados en la generación de nuevo conocimiento útil para la comunidad; sin embargo, existen ciertas limitaciones al momento de evaluar y corroborar los resultados obtenidos, dado que en muchos casos las publicaciones no exponen toda la información y los procesos que se emplearon durante la investigación (Fomel, 2015). Es por este motivo que es necesario establecer modelos, herramientas o metodologías que permitan la replicación de investigaciones en el campo de Ingeniería de Software para que el conocimiento obtenido no se mantenga aislado y pueda ser aceptado o rechazado por la comunidad científica. Adicionalmente (Sjøberg et al., 2005) mencionan que dentro de la comunidad de ingeniería de software existe la necesidad de que se realicen más estudios empíricos para elaborar o mejorar procesos, métodos y herramientas dedicados al desarrollo y mantenimiento de software

La replicación de investigaciones científicas tiene el potencial de servir como un estándar mínimo para juzgar afirmaciones científicas, aumentar la confiabilidad y confirmar teorías formuladas por los investigadores (Fomel, 2015). En este contexto, la intención de replicar las investigaciones por otros científicos ha dado lugar a lo que se conoce como investigación reproducible (Morgan, 2014). La presente investigación proporcionará una línea base de cómo aplicar el paradigma de la Investigación Reproducible a cierto tipo de estudios de Ingeniería de Software (encuestas), permitiendo que aquellas investigaciones que surjan a partir de esta línea base se pueden reproducir fácilmente por la comunidad científica, evitando que los resultados obtenidos se mantengan aislados y no se afiance el conocimiento científico de la Ingeniería de Software.

1.2 OBJETIVOS

1.2.1 Objetivo General

Proponer un modelo conceptual que permita la incorporación de la investigación reproducible en Ingeniería de Software, utilizando como estudio de caso el Proyecto NaPiRE de Ecuador desarrollado por el grupo GrIMPSoftII.

1.2.2 Objetivos Específicos

- Realizar un estudio de literatura inicial para identificar los métodos, técnicas y tecnologías que incluye la investigación reproducible en distintas disciplinas, con el propósito de adaptar este paradigma a la ingeniería de software.
- Plantear un modelo conceptual y de proceso para integrar la investigación reproducible en la ingeniería de software.

- Validar el modelo planteado a través del desarrollo de un prototipo que muestre la aplicabilidad del paradigma de investigación reproducible en un estudio de caso, en un grupo de investigación en ingeniería de software.

1.3 ALCANCE

Esta investigación comprende el planteamiento de un modelo conceptual para aplicar la investigación reproducible a estudios en ingeniería de software. Como estudio de caso se utilizará al proyecto NaPiRE, una investigación de tipo survey (investigación basada en encuestas), desarrollado por el grupo GrIMPSoftII, el cual consiste en un conjunto de entrevistas distribuidas globalmente acerca de prácticas y problemas en la Ingeniería de Requerimientos (Méndez Fernández et al., 2015). NaPiRE es una iniciativa académica que pretende establecer una teoría holística acerca de las prácticas y problemas industriales en la Ingeniería de Requerimientos (Méndez, Wagner, Felderer, & Kalinowski, 2012). Con el fin de validar el modelo propuesto, se desarrollará un prototipo de plataforma informática que permita a los investigadores reproducir el caso de estudio.

Para delinear de forma adecuada el alcance de la investigación planteada, se proponen varias preguntas de investigación asociadas a los objetivos específicos (ver Tabla 1).

Tabla 1
Preguntas de Investigación

| OBJETIVO ESPECÍFICO | PREGUNTA DE INVESTIGACIÓN |
|---|---|
| i. Realizar un estudio de literatura inicial para identificar los métodos, técnicas y tecnologías que incluye la investigación reproducible en distintas disciplinas, con el propósito de adaptar este paradigma a la ingeniería de software. | <p>RQ1. ¿Qué métodos se ha utilizado para aplicar el paradigma de investigación reproducible en las distintas disciplinas?</p> <p>RQ2. ¿Qué tecnología se ha utilizado para aplicar el paradigma de investigación</p> |

CONTINÚA 

| | |
|---|---|
| | reproducible en las distintas disciplinas? |
| ii. Plantear un modelo conceptual y de proceso para integrar la investigación reproducible en la ingeniería de software. | RQ3. ¿Cuáles son los beneficios obtenidos al integrar in investigación reproducible en la ingeniería de software? |
| | RQ4. ¿Qué tecnologías debe contemplar el modelo conceptual que integre investigación reproducible en la ingeniería de software? |
| iii. Validar el modelo planteado a través del desarrollo de un prototipo que muestre la aplicabilidad del paradigma de investigación reproducible en un estudio de caso, en un grupo de investigación en ingeniería de software. | RQ5. ¿Es posible desarrollar una herramienta que ayude a aplicar investigación reproducible en estudios de ingeniería de software? |
| | RQ6. ¿Qué arquitectura de software es la más óptima para desarrollar el prototipo que aplique el paradigma de investigación reproducible en el estudio de caso seleccionado? |

1.4 HIPÓTESIS

El paradigma de investigación reproducible adoptada de otras disciplinas servirá como modelo y facilitará la replicación de estudios en Ingeniería de Software.

CAPÍTULO II

MARCO METODOLÓGICO

2.1. MARCO TEÓRICO.

A continuación, se indican las variables que conforman la hipótesis propuesta, misma que se encuentra en el capítulo anterior.

2.1.1. Señalamiento de variables.

- **Variable independiente:** paradigma de investigación reproducible.
- **Variable dependiente:** facilidad de reproducción de estudios en ingeniería de software.

2.1.1.1.Red de categorías.

Con el objetivo de fundamentar el marco teórico y obtener sustento científico para el presente estudio, se estableció una red de categorías basada en las variables que interactúan en la investigación. El esquema de red de categoría se detalla en la Figura 1.

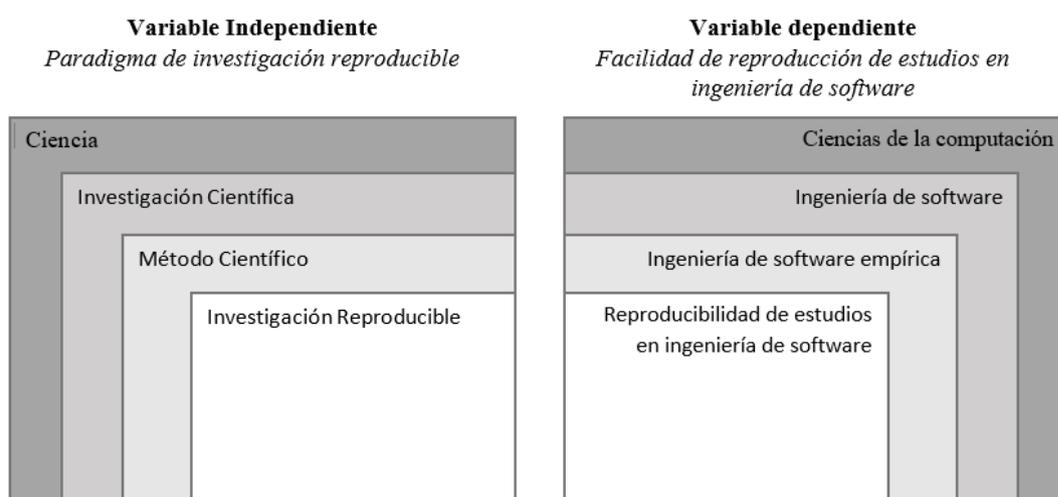


Figura 1. Red de categorías de las variables de investigación

2.1.2. Ciencia

La ciencia surge de la actividad conjunta de los individuos en la sociedad, se resume en conocimiento del mundo que los rodea y nace como resultado del trabajo intelectual de los hombres. La ciencia conforma un sistema de conocimientos como resultado del proceso de investigación científica, este sistema de conocimientos contribuye a la solución de los problemas que enfrenta el hombre en su relación con su medio a través de: principios, categorías, leyes y teorías, que son el contenido fundamental de toda ciencia (Álvarez de Zayas & Sierra, 1998).

El proceso de descubrimiento científico no está limitado a científicos profesionales que trabajan en laboratorios. La experiencia diaria de deducir qué está sucediendo con el automóvil que no enciende comparte similitudes con los descubrimientos en la ciencia. Dichas actividades involucran realizar observaciones, analizar evidencia, comprobar ideas y aferrarse a aquellas que sí funcionan (Undersea, Northwest, & Bauer, 2013).

Los cambios tecnológicos en la actualidad surgen gracias a los aportes que genera la ciencia dentro del entorno socio-cultural de la humanidad (Álvarez de Zayas & Sierra, 1998). En dicho contexto se estudia el mundo natural, esto incluye los componentes del universo físico que rodea a los seres humanos, como son: átomos, plantas, ecosistemas, la población e incluso universos (Undersea et al., 2013). Toda ciencia busca encontrar explicaciones a hechos en el mundo real y la sustenta con evidencia con el fin de obtener beneficio para la comunidad científica así como también para toda la humanidad que pueda aprovechar este nuevo conocimiento (Undersea et al., 2013).

2.1.3. Investigación Científica

La investigación surge como necesidad de búsqueda de conocimiento, dicho conocimiento implica una búsqueda científica y sistemática de información pertinente sobre un tema específico sobre cualquier rama de conocimiento (Álvarez de Zayas & Sierra, 1998). Redman y Mory definen la investigación como un esfuerzo sistematizado para ganar nuevo conocimiento un conjunto de acciones planificadas que se emprenden con la finalidad de resolver, total o parcialmente, un problema científico determinado. Además, la investigación implica que el investigador está interesado en más de un resultado particular, él está interesado en la reproducibilidad de sus resultados y en su extensión a situaciones más complejas y generales (Kothari, 2004).

2.1.3.1. Objetivos de la investigación.

El propósito de la investigación científica es descubrir respuestas a preguntas a través de la aplicación de procedimientos científicos para la creación de nuevos conocimientos o tecnologías (Álvarez de Zayas & Sierra, 1998). El enfoque principal de la investigación es encontrar la verdad que ha estado escondida y que aún no ha sido explorada. Al observar que cada investigación tiene un propósito concreto, se puede pensar que dichos objetivos de investigación se encasillan en uno de los siguientes grupos (Kothari, 2004).

- Con el fin de ganar familiaridad con un fenómeno o para lograr nuevos conocimientos sobre este (estos estudios también se denominan exploratorios)(Álvarez de Zayas & Sierra, 1998).
- Con el fin de retratar con precisión las características de un individuo en particular, una situación o un grupo (también se denomina estudios descriptivos)(Álvarez de Zayas & Sierra, 1998).
- Con el fin de determinar la frecuencia con la que ocurre algo o con lo que este

se asocia (se los puede llamar también estudios de diagnóstico)(Kothari, 2004).

- Con el fin de comprobar una hipótesis o una relación causal entre variables (también denominados estudios de hipótesis)(Álvarez de Zayas & Sierra, 1998).

2.1.3.2.Características de la Investigación Científica.

La investigación científica es sistemática ya tiene procedimientos estandarizados, de manera que se rige a una manera constante de hacer las cosas. Se la realiza de manera controlada y no deja hechos al azar, suele ser comparativa. Es empírica ya que es aplicable a fenómenos observables y medibles de la realidad. Tiene un enfoque racional-crítico dejando de lado los juicios de valor y analizando todo de forma objetiva. Pretende ser reproducible al implementar procedimientos estandarizados y sistemáticos (Sáenz Campos & Tinoco Mora, 1999).

2.1.3.3.Tipos de Investigación Científica.

Se van a exponer los tipos de investigación a modo de comparación entre semejantes. Los tipos básicos de investigación son los siguientes.

Descriptivo vs. Analítico.

La investigación descriptiva incluye encuestas y consultas de investigación de diferentes tipos. El objetivo principal de esta investigación es describir la situación tal como existe en la actualidad (Sáenz Campos & Tinoco Mora, 1999). La principal característica de este método es que el investigador no tiene control sobre las variables; él solo puede informar lo que ha sucedido o lo que está sucediendo. Estos estudios también incluyen los intentos de los investigadores para descubrir las causas incluso cuando no pueden controlar las variables. Los métodos de investigación utilizados en la investigación descriptiva son métodos de encuesta de todo tipo, incluidos los métodos comparativos y correlacionales. En la investigación analítica, por otro lado, el investigador debe utilizar hechos o información ya disponible, y analizarlos para hacer una evaluación crítica del material (Kothari, 2004).

Aplicado vs. Fundamental.

La investigación puede ser aplicada (o acción) o investigación fundamental (básica o pura). La investigación aplicada tiene como objetivo encontrar una solución para un problema inmediato al que se enfrenta una sociedad o una organización industrial / empresarial, mientras que la investigación fundamental se ocupa principalmente de las generalizaciones y de la formulación de una teoría (Hernández Sampieri, Fernández Collado, & Baptista Lucio, 2014). Reunir el conocimiento por el conocimiento se denomina investigación 'pura' o 'básica'. Investigar acerca de algún fenómeno natural o relacionado con las matemáticas puras son ejemplos de investigación fundamental, pero la investigación dirigida a ciertas conclusiones (por ejemplo, una solución) frente a un problema social o empresarial concreto es un ejemplo de investigación aplicada. Por lo tanto, el objetivo central de la investigación aplicada es descubrir una solución para un problema práctico apremiante, mientras que la investigación básica se dirige a encontrar información que tiene una amplia base de aplicaciones y, por lo tanto, se suma al cuerpo organizado ya existente de conocimiento científico (Kothari, 2004).

Cuantitativo vs. Cualitativo.

La investigación cuantitativa se basa en la medición de cantidad o cantidad. Es aplicable a fenómenos que se pueden expresar en términos de cantidad (Hernández Sampieri et al., 2014). La investigación cualitativa, por otro lado, se ocupa de los fenómenos cualitativos, es decir, los fenómenos relacionados con o que involucran calidad o bondad (Kothari, 2004). Por ejemplo, cuando estamos interesados en investigar las razones del comportamiento humano (es decir, por qué las personas piensan o hacen ciertas cosas), con frecuencia hablamos de "Investigación de Motivación", un tipo importante de investigación cualitativa. La investigación de actitud o de opinión, es decir, la investigación diseñada para descubrir cómo se sienten las personas o lo que piensan sobre un tema o institución en particular también es una investigación cualitativa.

La investigación cualitativa es especialmente importante en las ciencias del comportamiento, donde el objetivo es descubrir los motivos subyacentes del comportamiento humano (Kothari, 2004).

Conceptual vs. Empírico.

La investigación conceptual es aquella relacionada con alguna idea o ideas abstractas. Generalmente es utilizado por filósofos y pensadores para desarrollar nuevos conceptos o para reinterpretar los existentes (Hernández León & Coello Gonzáles, 2011). Por otro lado, la investigación empírica depende solo de la experiencia o la observación, a menudo sin la debida consideración del sistema y la teoría. Es una investigación basada en datos que presenta conclusiones que pueden verificarse mediante observación o experimento. También podemos llamarlo como tipo de investigación experimental. En una investigación de este tipo es necesario conocer los hechos de primera mano, en su origen, y hacer activamente ciertas cosas para estimular la producción de la información deseada. En una investigación de este tipo, el investigador primero debe proporcionarse una hipótesis de trabajo o adivinar los resultados probables. Luego trabaja para obtener suficientes datos (datos) para probar o refutar su hipótesis. A continuación, establece diseños experimentales que cree que manipularán a las personas o los materiales en cuestión a fin de generar la información deseada. Tal investigación se caracteriza así por el control del experimentador sobre las variables en estudio y su manipulación deliberada de una de ellas para estudiar sus efectos. La investigación empírica es apropiada cuando se busca la prueba de que ciertas variables afectan de algún modo a otras variables. La evidencia reunida a través de experimentos o estudios empíricos hoy se considera el soporte más poderoso posible para una hipótesis dada (Kothari, 2004).

2.1.4. Método Científico

El método científico es uno solo, y el mismo en las ramas (de la ciencia) a la vez que es el método de todas las mentes lógicamente entrenadas. La unidad de todas las ciencias consiste solo en sus métodos, no en su material; el hombre que clasifica hechos de cualquier tipo, que ve su relación mutua y describe sus secuencias, está aplicando el Método Científico y es un hombre de ciencia (Pearson, 1999).

Visto de manera rápida, la mayoría de los libros explican y agrupan al método científico en las siguientes fases: formulación de la hipótesis, prueba de la hipótesis, lógica deductiva e inductiva, experimentos controlados, interacción entre datos y teoría, límites del dominio de ciencia. Los principios generales del método científico son de tres tipos: principios que son relativamente distintivos de la ciencia misma, principios más amplios que se encuentran en todas las formas de investigación racional y principios básicos con sus fuentes en sentido común. (Gauch, 2003).

2.1.5. Investigación Reproducible

La investigación reproducible, según la definición de Jon Claerbout (Claerbout, 1991), se refiere a la disciplina de adjuntar códigos y datos de software a publicaciones científicas, a fin de permitir la verificación y replicación independientes de los experimentos. El llamado “principio de Claerbout” afirma que “Un artículo sobre ciencia computacional en una publicación científica no es el conocimiento en sí, es simplemente una publicidad del conocimiento. El conocimiento actual es el entorno de desarrollo de software completo y el conjunto completo de instrucciones que generaron las cifras” (Fomel et al., 2013). Además de ayudar a los autores a reproducir su propio trabajo, compartir la investigación con otros también abre muchas posibilidades para la colaboración en la investigación que, de lo contrario, se cerraría (Fomel, 2015).

Generalmente existen dos formas de replicación, interna y externa. Replicación interna es llevada a cabo por el investigador o grupo de investigadores originales, mientras que la externa la realizan investigadores ajenos al grupo de investigadores originales. Además, el grado en que se lleva a cabo una replicación puede dividirse en replicación exacta y conceptual. Una replicación exacta es una replicación que sigue el procedimiento original lo más cercano posible, mientras que la replicación conceptual es una replicación donde se valida la misma hipótesis mediante un procedimiento diferente (Larsson et al., 2014).

La investigación reproducible busca como objetivo la reproducibilidad computacional, es decir, la capacidad de un segundo investigador para recibir un conjunto de archivos, incluidos datos, códigos y documentación, y para recrear o recuperar los resultados de un proyecto de investigación, incluidas figuras, tablas, y otros resultados cuantitativos y cualitativos clave (Kitzes, Turek, & Deniz, 2018). Para hacer que este estudio sea reproducible, todos los resultados se generan automáticamente y se pueden recrear en cualquier momento descargando los recursos del estudio (Sommerville, 2011) y ejecutando el script de análisis como se describe en la documentación adjunta (se necesitan R y varios paquetes)(Larsson et al., 2014).

Mientras los estudios empíricos se hacen más comunes en la Ingeniería de Software, la importancia de que estos sean reproducibles aumenta. Estas repeticiones son importantes ya que ayudan a aumentar el cuerpo de conocimiento alrededor de Ingeniería de Software, lo que a su vez conduce a una mayor madurez del campo. Una replicación de un estudio también viene con beneficios para el estudio original en términos de aumento de la confianza para el experimento conducido y los resultados informados (por ejemplo, los intervalos de confianza más estrictos). Además, el éxito de una replicación conducida no depende principalmente de cuán bien los resultados replicados se ajusten al original, sino de la contribución al conjunto de conocimientos (Larsson et al., 2014). Es necesario tomar en cuenta que no es posible sacar

conclusiones generales a partir de una o pocas repeticiones, sino que, se necesita consolidar los hallazgos en base a múltiples repeticiones que posteriormente sean combinadas (Fonseca Carrera, 2012).

2.1.5.1. Etapas básicas en Investigación Reproducible.

El flujo de trabajo de investigación reproducible básico se puede dividir en tres etapas principales: adquisición de datos, procesamiento de datos y análisis de datos. Estas tres etapas están precedidas por actividades relacionadas con la configuración del sistema, y son seguidas por pasos que automatizan todo el flujo de trabajo tanto como sea posible (Kitzes et al., 2018).

Adquisición de datos

La primera etapa del flujo de trabajo básico es la adquisición, entrada o creación de datos. Esta etapa generalmente consiste en recopilar datos de una fuente principal, como la observación de campo, la investigación experimental o las encuestas. Sin embargo, también puede incluir la adquisición de datos de una fuente existente, a través del raspado web o la comunicación con otros investigadores, o la generación de datos a través de la simulación. Independientemente del método, el resultado final de esta primera etapa son los datos sin procesar (Kitzes et al., 2018).

Procesamiento de datos

La segunda etapa implica el procesamiento o la limpieza de los datos producidos en la primera etapa. Dependiendo de las herramientas utilizadas y las estrategias del autor, esta etapa puede incluir tareas como la entrada manual de datos, la revisión visual de los datos o la manipulación o el filtrado sistemático de los datos utilizando scripts u otro software. Al finalizar esta segunda etapa, los datos relevantes se digitalizan, se limpian y se preparan completamente para el análisis (Kitzes et al., 2018).

Análisis de datos

La tercera etapa es el análisis de datos. La forma más común de análisis de datos es la estadística formal, pero las actividades en esta etapa incluyen la visualización de datos, la evaluación del rendimiento de algoritmos particulares y la extensión de los datos para abordar una hipótesis o extraer una conclusión científica. El atributo definitorio de esta etapa es que analiza, de alguna manera, los datos limpios producidos en la segunda etapa, y produce los productos científicos deseados de la investigación, generalmente resultados cuantitativos en forma de figuras y tablas que están incorporados en manuscritos, charlas, y otras formas de comunicación (Kitzes et al., 2018).

Finalmente, después de las tres etapas centrales, la reproducibilidad de un proyecto se puede mejorar enormemente mediante la creación de un solo script controlador que puede ejecutar automáticamente las tres etapas para producir un resultado final. Cuando este tipo de flujo de trabajo de "pulsador" es poco realista o imposible de lograr debido a las restricciones del proyecto, se debe crear una documentación detallada de todos los pasos no automatizados (Kitzes et al., 2018).

2.1.6. Ciencias De La Computación

Las ciencias de la computación son el estudio sistemático de procesos algorítmicos (teoría, análisis, diseño eficiencia, implementación y aplicación) que describen y transforman información; nació a inicios de los años 40s bajo la pregunta fundamental ¿Qué puede ser eficientemente automatizado?, juntando la teoría de algoritmos, lógica matemática y con la intervención de la computadora electrónica con programas almacenados (Chairman et al., 1989).

Las ciencias de la computación se encuentran interrelacionada profundamente con las matemáticas quienes proporcionan los análisis en el campo y la ingeniería quien proporcionan

los diseños; esta disciplina acoge su propia teoría, método experimental e ingeniería lo que la distingue de la mayoría de ciencias físicas puesto que se encuentran separadas de las ingenierías que aplican sus hallazgos (Chairman et al., 1989).

2.1.6.1.Algoritmos y estructuras de datos

Esta área se encarga del estudio de los algoritmos en aspectos como eficiencia, comportamiento y recursos utilizados; adicionalmente se centra en encontrar la solución más eficiente a las diferentes clases de problemas utilizando algoritmos y determina si estos pueden ser aplicados a problemas similares; finalmente estudia la mejor forma de almacenar y acceder a la información (Chairman et al., 1989).

2.1.6.2.Lenguajes de programación

Se encarga del estudio de notaciones para máquinas virtuales que ejecutan algoritmos, la notación de los algoritmos y la información y finalmente la traducción eficiente de lenguajes de alto nivel a códigos de máquina (Chairman et al., 1989).

2.1.6.3.Arquitectura

Se encarga del estudio de los métodos para organizar el hardware y su software asociado en sistemas eficientes y confiables, adicionalmente se centra en el estudio de aspectos como el diseño, control, comportamiento y seguridad de estos conjuntos de hardware (Chairman et al., 1989).

2.1.6.4.Computación numérica y simbólica

La computación numérica y simbólica se encarga del estudio de métodos para resolver ecuaciones de modelos matemáticos de sistemas informáticos de forma eficiente y precisa (Chairman et al., 1989).

2.1.6.5.Sistemas Operativas

Los sistemas operativos tienen como área de interés para su estudio todos aquellos mecanismos de control que permiten a múltiples recursos (software y hardware) coordinarse eficientemente durante la ejecución de programas (Chairman et al., 1989).

2.1.6.6.Metodología e ingeniería de software

La metodología e ingeniería de software tiene como área de estudio los métodos, técnicas y procesos para el diseño de programas y grandes sistemas de software que cumplen con sus especificaciones y son prudentes, seguros, confiables y dependientes (Chairman et al., 1989).

2.1.6.7.Bases de datos y recuperación de información

El área de base de datos y recuperación de información tiene como objeto de estudio los métodos, técnicas y procesos de organización de grandes conjuntos de información persistente y compartida para su eficiente almacenamiento y actualización (Chairman et al., 1989).

2.1.6.8.Inteligencia artificial y robótica

El área de inteligencia artificial y robótica se encarga del estudio de los procesos, técnicas y métodos para modelar, replicar o emular el comportamiento del ser humano y de los animales (Chairman et al., 1989).

2.1.6.9.Comunicación humano computadora

El área de comunicación humano computadora tiene como objeto de estudio el conjunto de prácticas, procesos y métodos que definen los medios eficientes para transferir información entre los humanos y las máquinas mediante el uso de sensores y motores que reflejan los cálculos humanos (Chairman et al., 1989).

2.1.7. Ingeniería de Software

La ingeniería de software es la creación u mantenimiento de software. Sin embargo,

desde una perspectiva de investigación la ingeniería de software es la base de conocimiento acerca de la creación y mantenimiento de software; de la misma forma es base de conocimiento acerca del fenómeno emergente de estas dos actividades (Monperrus, 2016).

La ingeniería de software es una disciplina de las ciencias de computación que ofrece métodos y técnicas para desarrollar y mantener software de calidad que sirven para resolver problemas de todo tipo. La ingeniería de software desde la perspectiva de Zilkovitz es el estudio de los principios y metodologías para desarrollo y mantenimiento de sistemas (Pressman, 2010).

Según la IEEE la ingeniería de software es la aplicación de un enfoque sistemático, disciplinado y cuantificable al desarrollo, operación y mantenimiento de software, es decir la aplicación de la ingeniería de software como tal y así mismo se la define como el estudio de los enfoques anteriormente expuestos (Pressman, 2010).

2.1.8. Ingeniería de Software Empírica

La ingeniería de software empírica es un área de investigación interesada en observaciones empíricas de artefactos de ingeniería de software y validaciones empíricas de teorías y suposiciones de la ingeniería de software (Monperrus, 2016).

2.1.8.1. Experimento controlado

Es una investigación de hipótesis comprobable donde una o más variables independientes con manipuladas para medir su efecto en uno o más variables dependiente; los experimentos controlados permiten determinar de forma precisa la relación de causa efecto que existe entre las variables independientes con las dependientes (Easterbrook, Singer, Storey, & Damian, 2008).

2.1.8.2.Caso de estudio

Erróneamente se entiende como caso de estudio a un ejemplo funcional, como un método empírico significa otra cosa; Yin presenta el caso de estudio como “una pregunta empírica que investiga a un fenómeno contemporáneo dentro de su contexto real, especialmente cuando los límites entre el fenómeno y su contexto no son evidentemente claros”, los casos de estudio ofrecen un entendimiento profundo del cómo y él por que ocurren ciertos fenómenos y permiten revelar el mecanismo por el cual las relaciones de causa y efecto ocurren (Easterbrook et al., 2008).

2.1.8.3.Encuesta

La investigación con encuestas es utilizada para identificar las características de una amplia población de individuos. Está asociada cercanamente con el uso de cuestionarios para la recolección de información; sin embargo se puede conducir utilizando entrevistas estructuradas o técnicas de bitácoras de información. La característica fundamental de la investigación con encuestas es la selección de la muestra representativa de una población bien definida y las técnicas de análisis de información usadas para generalizar desde la muestra hasta la población (Easterbrook et al., 2008).

2.1.8.4.Etnografía

Es una forma de investigación que se enfoca en la sociología del significado mediante la observación de campo. De acuerdo con Robinson el objetivo es estudiar a una comunidad de personas para entender como sus miembros tienen sentido en sus interacciones sociales. En ingeniería de software etnografía ayuda a entender como las comunidades tecnológicas construyen una cultura de prácticas y estrategias de comunicación que les permite realizar trabajo técnico de manera colectiva (Easterbrook et al., 2008).

2.1.8.5. Investigación de acción

Davison afirma que en la investigación de acción los investigadores intentan resolver problemas del mundo real mientras que simultáneamente estudian la experiencia de resolver el problema. Mientras la mayoría de estudios empíricos intentan observar el mundo como existe actualmente, los investigadores de acción apuntan a intervenir en las situaciones estudiadas con el propósito de mejorar la situación (Easterbrook et al., 2008).

2.2. ESTADO DEL ARTE

Para analizar el estado del arte acerca del uso de la investigación reproducible en la ingeniería de software se llevó a cabo una revisión de literatura inicial basada en las guías de revisión sistemática de literatura propuestas por (Kitchenham & Charters, 2007) (ver Figura 2).

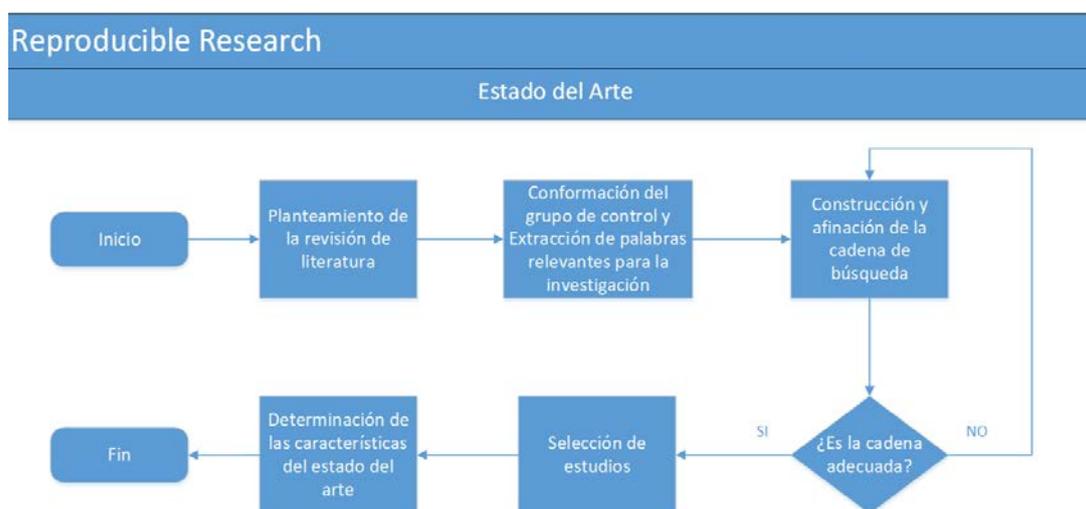


Figura 2. Método propuesto para la elaboración del estado del arte

2.2.1. Planteamiento de la revisión de literatura

En esta fase se realizó una breve descripción del problema de investigación para proporcionar un contexto para la búsqueda de estudios científicos; posteriormente se procedió a definir un objetivo de búsqueda y plantear preguntas de investigación para alinear la búsqueda en relación al problema de investigación y finalmente se definieron los criterios de inclusión y exclusión.

2.2.2. Conformación del grupo de control (GC)

Según (Petersen, Feldt, Mujtaba, & Mattsson, 2008) un paso fundamental de un mapeo sistemático de literatura es definir o delimitar los artículos que se consideren relevantes para la investigación, eliminando aquellos que solo mencionan nuestro enfoque principal.

Fue necesaria la participación de al menos tres investigadores. Cada investigador propuso estudios que podrían ser parte del grupo de control. Posteriormente se realizó una validación cruzada donde se estableció finalmente el grupo de control (ver Tabla 2).

Tabla 2

Grupo de Control

| <i>TÍTULO</i> | <i>CITA</i> | <i>PALABRAS CLAVE</i> |
|--|--|--|
| Outliers and Replication in Software Engineering | (Larsson et al., 2014) | software engineering, outliers, replication, reproducibility. |
| Madagascar open-source software project for multidimensional data analysis and reproducible computational experiments. | (Fomel et al., 2013) | reproducibility, data analysis, geophysics, seismology, Python |
| Conducting Reproducible Research with Umbrella | (Meng, Thain, Vyushkov, Wolf, & Woodard, 2016) | replication, reproducibility, software, software engineering |

Tras un análisis de los estudios del GC, se seleccionaron las palabras más relevantes respecto al objetivo de la búsqueda, en este caso fueron: reproducibility, software engineering, replication, scientific computing, scientific software, reproducible research.

2.2.3. Construcción y afinación de la cadena de búsqueda

Con las palabras clave que fueron obtenidas de los artículos científicos del grupo de control se conformó la cadena de búsqueda: (“SOFTWARE ENGINEERING” OR “SOFTWARE” OR “COMPUTATIONAL SCIENCE” OR “EMPIRICAL SOFTWARE ENGINEERING”) AND (“REPRODUCIBLE RESEARCH” OR “REPLICATION” OR “REPRODUCIBILITY” OR “FAMILY OF EXPERIMENTS”) AND (“FRAMEWORK” OR

“SOFTWARE” OR “OPEN-SOURCE” OR “TOOLS”), misma que se utilizó en la base digital IEEE Explore.

Sin embargo, esta cadena obtuvo un gran número de investigaciones inclusive aquellas que fueron descartadas. Después de realizar varias pruebas con distintas combinaciones de cadenas, se seleccionó la cadena: “(((SOFTWARE ENGINEERING, COMPUTER SCIENCE) AND REPRODUCIBLE RESEARCH, REPLICATION) AND TOOL, FRAMEWORK)”.

2.2.4. Selección de estudios

Al aplicar la cadena de búsqueda en la base digital IEEE Explore se obtuvo alrededor de 80 artículos relacionados con el tema el cual se consideró un número de artículos manejable; adicionalmente con esta cadena la mayor parte de los artículos del grupo de control apareció dentro de los artículos encontrados.

De los 80 artículos obtenidos se aplicaron dos filtros, los cuales son explicados a continuación:

- **Vigencia:** Estudios realizados a partir del año 2011. Se eligió este año debido a que la tecnología avanza rápidamente, por lo que es necesario tener estudios con una relativa actualidad.
- **Tipo de estudio:** Fueron elegidos únicamente estudios del tipo: technical report, conference paper y journal paper; debido a su relevancia.

En base a los filtros antes mencionados, y después de analizar si el contenido de los estudios aporta a responder las preguntas de investigación, se eligieron 6 estudios primarios, los cuales constituyen la base para realizar el estudio del estado del arte (ver Tabla 3).

Tabla 3*Estudio Seleccionados*

| CÓDIGO | TÍTULO | CITA |
|---------------|--|--|
| EP1 | Reproducible Research in Computational Science | (Peng, 2011) |
| EP2 | Conducting reproducible research with Umbrella: Tracking, creating, and preserving execution environments | (Meng et al., 2016) |
| EP3 | A Multi-Language Computing Environment for Literate Programming and Reproducible Research | (Schulte, Davison, Dye, & Dominik, 2012) |
| EP4 | Madagascar: open-source software project for multidimensional data analysis and reproducible computational experiments | (Fomel et al., 2013) |
| EP5 | Reproducible Research as a Community Effort: Lessons from the Madagascar Project | (Fomel, 2015) |
| EP6 | Outliers and replication in software engineering | (Larsson et al., 2014) |

2.2.5. Determinación de las características del estado del arte.

2.2.5.1. Resumen de los estudios primarios.

EP1 (Peng, 2011): Reproducible Research in Computational Science

El paper “Reproducible Research in Computational Science” aborda de forma general los lineamientos bases para poder aplicar investigación reproducible en investigaciones, adicionalmente dice que la reproducibilidad de un estudio en las ciencias de la computación consiste en poner a disposición de otros investigadores la información, técnicas y herramientas que se utilizaron para procesar dicha información, en esta investigación propone una forma de compartir los datos y el software utilizado por parte de los investigadores para que su estudio se pueda recrear por cualquier otro investigador. (Madrid & Macchetto, 2009). Sin embargo, esta aproximación carece de un estado de reproducibilidad total, debido a que se analiza la misma información del estudio original en repetidas ocasiones por distintos investigadores y carece de la posibilidad de aplicar el estudio en distintas condiciones para obtener un contraste

de la información obtenida.

EP2 (Meng et al., 2016): Conducting reproducible research with Umbrella: Tracking, creating, and preserving execution environments

En el paper “Conducting reproducible research with Umbrella: Tracking, creating, and preserving execution environments” sus autores hablan acerca de las dificultades que encuentran los investigadores interesados en replicar el trabajo de otros, puesto que en la mayoría de los casos no se detalla el procedimiento que existió durante el proyecto Umbrella brinda un framework para la investigación reproducible de experimentos computacionales, su enfoque consiste en capturar todas las configuraciones, especificaciones y dependencias del ambiente de experimentación utilizado para la investigación y posteriormente mediante el uso de herramientas como Docker o AmazonWS crear instancias de dicho ambiente para que otro investigador pueda replicar o realizar sus propios experimentos, actualmente dicho framework se enfoca en la obtención de resultados, mas no en el rendimiento del ambiente y trabaja exclusivamente con sistemas no distribuidos (Meng, Thain, Vyushkov, Wolf, & Woodard, 2017). Esta propuesta es un gran apoyo para reproducir investigaciones, solo si estas dependen de un ambiente pequeño que no cuente con muchos requerimientos de hardware, puesto que los costos para desplegar instancias con grandes demandas de hardware son elevados y esto dificulta que las investigaciones se reproduzcan.

EP3 (Schulte et al., 2012): A Multi-Language Computing Environment for Literate Programming and Reproducible Research

Eric Schulte en su paper “A Multi-Language Computing Environment for Literate Programming and Reproducible Research” explica la necesidad de unir el lenguaje natural con el lenguaje computacional para realizar documentos científicos, se enfoca en su propuesta el proyecto Org-mode proporciona un conjunto de herramientas para la programación de literatura

y la investigación reproducible debido a que permiten combinar la información, el código fuente y el texto de un proyecto en un solo documento; entre las ventajas del Org-mode es que puede combinar código fuente de diferentes lenguajes de programación y adicionalmente los archivos generados por este framework se puede exportar a una variedad de formatos (Schulte et al., 2012). El proyecto Org-mode es de gran ayuda para reproducir los resultados de una investigación de manera dinámica y plasmarlos en un documento formal que facilita al investigador interesado publicar un análisis de los datos analizados en la investigación para conocimiento de la comunidad científica.

EP4 (Fomel et al., 2013): Madagascar: open-source software project for multidimensional data analysis and reproducible computational experiments

Sergey Fomel et al. En su paper “Madagascar: open-source software project for multidimensional data analysis and reproducible computational experiments” explica cómo funciona la herramienta Madagascar y algunos resultados que se han tenido como consecuencia de su uso en la comunidad de investigadores enfocado a la geofísica. Madagascar Project está compuesto por tres niveles: módulos de línea de comandos (C, C++, Fortran-77, Matlab, Python), workflow aplicado a datos (Extended Python), publicación (LaTeX, Python). La principal meta de Madagascar Project es permitir la reutilización de “recetas computacionales” de estudios previos a los investigadores para que puedan generar nuevos resultados de investigación sobre los ya existentes (Fomel et al., 2013). Evidentemente, el proyecto Madagascar no ha sido aplicado en la Ingeniería de Software, los resultados reportados indican que ha sido bastante aplicado en investigaciones geofísicas.

EP5 (Fomel, 2015): Reproducible Research as a Community Effort: Lessons from the Madagascar Project

El artículo de la revista “Scientific Software Communities” de la Universidad de Texas

en Austin escrito por Segey Fomel se centra en los aprendizajes obtenidos después de 10 años de trabajo de Madagascar Project, una herramienta para reproducir investigaciones. Madagascar Project es la colección de investigaciones y sus recursos abiertos para su libre uso, consta de tres niveles básicamente que son: papers, gráficos y figuras y scripts utilizados por los científicos para sus investigaciones. Dentro de las lecciones aprendidas el autor menciona que la reproducibilidad requiere mantenibilidad y la mantenibilidad requiere de una comunidad abierta, por otra parte la reproducibilidad de una investigación no es exactamente la meta, ya que existen investigaciones donde es imposible aplicar la reproducibilidad, de todas formas, es beneficioso tener los recursos a disposición porque ayudan al lector a verificar y ampliar el entendimiento sobre la investigación publicada (Fomel, 2015). Después de leer los resultados obtenidos tras varios años de trabajo con Madagascar Project se evidencia un gran repositorio de “recetas computacionales” que pueden servir para reproducir ciertas partes de investigaciones en diferentes campos de estudio, sin embargo, no existen replicas de investigaciones en su totalidad que comprenda todo el proceso investigativo específico para cierto tipo de investigación.

EP6 (Larsson et al., 2014): Outliers and replication in software engineering

Henrik Larsson et al. realizaron un estudio de Reproducible Research enfocado a los outliers (información de una investigación cuyos valores se desvían significativamente del resto) y la Ingeniería de Software. Los autores mencionan que la disponibilidad de información dentro de la Investigación en Ingeniería de Software es muy baja por lo que dificulta la replicación de los estudios. De los estudios que seleccionaron, solo el 37% tuvieron su información accesible. En cuanto al manejo de outliers, el 63% de investigaciones no mencionaron cómo fueron manejados dichos outliers (Larsson et al., 2014). Con estos datos, los autores ayudan a resaltar la importancia de realizar una investigación que busque mejorar

la reproducibilidad específicamente en Ingeniería de Software.

2.2.5.2. Características del estado del arte

Existe una cantidad limitada de estudios relevantes en cuanto a la aplicación de la investigación reproducible en estudios empíricos de Ingeniería de Software, la mayoría no tiene un campo de aplicación delimitado por lo que existen datos muy dispersos e incompletos. Por otro lado, la mayoría de herramientas existentes utilizan línea de comandos para su ejecución por lo que su usabilidad se ve afectada. Los estudios mencionados anteriormente brindan algunos métodos y formas de replicación que han sido utilizadas, donde se identificó que la principal limitación es que no se ha podido establecer alguna herramienta para realizar reproducibilidad total aplicada a un área específica. Esta dificultad abre la puerta para proponer un modelo que cubra un tipo de investigación específica, el cual facilite a otros investigadores la replicación del estudio.

2.3. METODOLOGÍA DE LA INVESTIGACIÓN

2.3.1. Investigación Exploratoria De Método Inductivo

La investigación es de carácter exploratorio, debido a que se centra en analizar e investigar aspectos concretos de la realidad que aún no han sido analizados a profundidad (Shields & Rangarajan, 2013). Así mismo, la investigación es de carácter inductiva, debido a que se basa en la obtención de conclusiones a partir de la observación de hechos (Goddard & Melville, 2001). Adicionalmente, la investigación es de carácter cuantitativo, dado que utiliza valores numéricos y métodos estadísticos para explicar ciertos aspectos involucrados en el problema analizado (Kothari, 2004).

Con el propósito de llevar a cabo un proceso riguroso de trabajo, se seleccionó la metodología Design Science (Wieringa, 2013) como guía del proyecto, la cual se describe a detalle en la Sección 2.3.

2.3.2. Metodología Design Science

Design science es una metodología aplicada comúnmente en proyectos referentes a las Ciencias de la Computación, y consiste en el diseño y la investigación de artefactos que pertenezcan al contexto. Los artefactos pueden ser de cuatro tipos, los cuales se describen a continuación:

- **Constructos:** proveen el lenguaje en el cual problemas y soluciones son definidos y comunicados. Puede ser vocabulario o símbolos (Cataldo, 2015).
- **Modelos:** representan el problema y solución y la conexión entre sus componentes. Ejemplo de este tipo son abstracciones y representaciones (Cataldo, 2015).
- **Métodos:** proveen guías sobre cómo resolver problemas y encontrar las soluciones. A este tipo corresponden algoritmos y prácticas (Cataldo, 2015).
- **Instanciaciones:** muestran que construcciones, modelos o métodos pueden ser implementados. A esta categoría corresponden sistemas prototipos e implementaciones de los mismos (Cataldo, 2015).

Dichos artefactos están diseñados para interactuar con un contexto problemático a fin de mejorar o solucionar algo en ese contexto. Un contexto puede ser: software, hardware, gente, organizaciones, servicios, métodos, técnicas, normas, procesos de negocio, objetivos organizaciones, etc. Lo que contribuye a la solución de un problema es la interacción entre el artefacto y el contexto del problema (Wieringa, 2014).

El objetivo de Design Science es desarrollar conocimiento, de modo que el profesional de una disciplina específica lo pueda utilizar para diseñar soluciones en su campo de estudio (Van Aken, 2005). Este conocimiento se alcanza mediante la construcción y aplicación de un artefacto diseñado para el dominio de la problemática tratada (Hevner, A. R., March, S. T.,

Park, J. & Ram, 2004).

2.3.3. Ciclo de diseño e ingeniería en Design Science

La metodología Design Science es de naturaleza iterativa y consta de cinco fases: investigación del problema, diseño de la solución, validación de la solución, implementación de la solución y evaluación de la implementación (Wieringa, 2013). Todas estas fases con sus respectivas preguntas de conocimiento y problemas de diseño están descritas en la Figura 3.

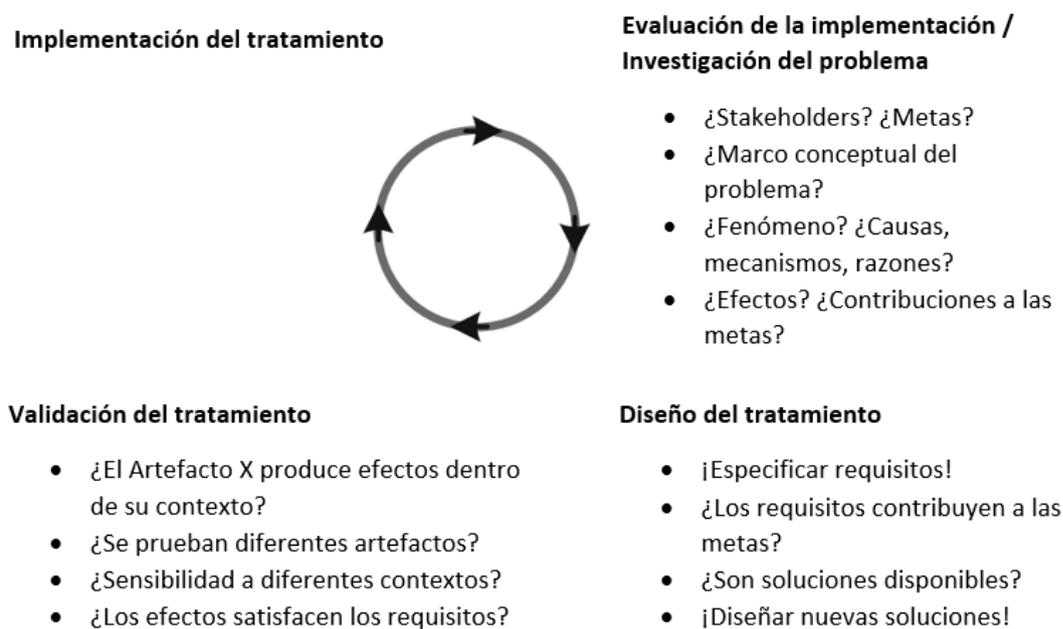


Figura 3. El ciclo de ingeniería

Fuente: (Wieringa, 2014)

2.3.3.1. Investigación del problema

En esta fase se entiende el problema y todos los elementos que lo conforman, de manera que se construye un marco conceptual alrededor del problema. En esta fase se interactúa con los Stakeholders de manera que se fijan metas para solucionar dicho problema (Wieringa, 2014). El problema debe tener relevancia, siendo importante su solución y que genere un beneficio para la comunidad (Cataldo, 2015).

Previo a definir el problema se llevó a cabo una revisión de literatura inicial basado en las guías de revisión sistemática de literatura propuestas por (Kitchenham & Charters, 2007).

Posteriormente se utilizó el árbol de problemas (Martinez & Fernández, 2015) como herramienta para profundizar e identificar el problema, causas y consecuencias involucrados en esta investigación.

2.3.3.2. Diseño del tratamiento

Es costumbre que los ingenieros digan que están diseñando soluciones, pero se evita este término porque puede impedir ver la posibilidad de que un artefacto pueda resolver un problema solo parcialmente o tal vez no. Los artefactos pueden incluso introducir nuevos problemas. Tratamiento es la interacción entre el artefacto y el contexto del problema. El investigador en Design Science no solo diseña un artefacto, sino que diseña una interacción entre el artefacto y el contexto del problema, enfocada en tratar el problema (Wieringa, 2014).

Los tratamientos, y por lo tanto los artefactos, están diseñados, y estos diseños están documentados en una especificación. Existe una considerable diversidad en el uso de estos términos en la ingeniería de software y otras ramas de la ingeniería. En ingeniería de software, una especificación describe el comportamiento externo del software. Un diseño es una decisión sobre qué hacer y una especificación es una documentación de esa decisión (Wieringa, 2014).

Los artefactos diseñados deben ser viables y de interés. Al momento de diseñarlo puede haber dudas sobre la posibilidad de que funcione, estas dudas serán resueltas en las siguientes fases (Cataldo, 2015).

Con el fin de buscar y definir los artefactos que se diseñaron y construyeron, se utilizó la técnica de lluvia de ideas (Zapata Domínguez, Murillo Vargas, & Martínez Crespo, 2009), misma que fue aplicada conjuntamente con el director de la tesis. Se definió construir un prototipo de plataforma web para investigación reproducible, la misma que está basada en los modelos previamente diseñados, estos modelos son: modelo de procesos, modelo conceptual y modelo de base de datos. La metodología de desarrollo aplicado al prototipo fue la metodología

SCRUM (Martel, 2014) cuyos detalles se encuentran en el siguiente capítulo.

2.3.3.3. Validación del tratamiento

Validar un tratamiento es justificar que contribuiría a los objetivos de las partes interesadas si se implementa. En el ciclo de ingeniería, la validación se realiza antes de la implementación. Consiste en investigar los efectos de la interacción entre un prototipo de un artefacto y un modelo del contexto del problema y compararlos con los requisitos del tratamiento. Sobre esta base, se desarrolla una teoría del diseño, que es una teoría de las propiedades del artefacto y su interacción con el contexto del problema. Usando esta teoría, se predice lo que sucedería si se implementara el artefacto, es decir, si se transfiriera al contexto del problema. El objetivo de la validación es predecir cómo un artefacto interactuará con su contexto, sin observar realmente un artefacto implementado en un contexto del mundo real. La investigación de validación es experimental y generalmente se realiza en el laboratorio. En la investigación de validación, exponemos un prototipo de artefacto a varios escenarios presentados por un modelo del contexto, para ver cómo responde. Los métodos de investigación más utilizados son el modelado, la simulación y las pruebas, métodos que se denominan "experimentos de mecanismo de caso único" (Wieringa, 2014).

Después de construir el prototipo de plataforma web dedicada a la investigación reproducible se la sometió a pruebas de caja blanca y caja negra (Desikan & Ramesh, 2008) con el fin de comprobar que satisface los requerimientos planteados previamente en esta fase.

2.3.3.4. Implementación del tratamiento

La implementación de un tratamiento, se la define como la aplicación del tratamiento al contexto del problema original. Entonces, en lugar de la palabra "implementación", siempre se puede leer "transferir al contexto del problema". Lo que cuenta como una implementación depende del contexto del problema que tenemos en mente. Con respecto a los problemas del

mundo real, los proyectos de ciencias del diseño siempre están restringidos a las primeras tres tareas del ciclo de ingeniería: investigación de problemas, diseño de tratamientos y validación de tratamientos. Llamamos a estas tres tareas el ciclo de diseño, porque los proyectos de investigación en ciencias de diseño generalmente se llevan a cabo Varias iteraciones a lo largo de este ciclo. La confusión sobre lo que constituye una implementación aparece en los proyectos de ciencia del diseño de la siguiente manera: Desde el punto de vista de un actor externo, el investigador realiza algunas iteraciones a lo largo del ciclo de diseño y nunca produce una implementación en el contexto del problema social del proyecto de investigación. Pero desde el punto de vista del investigador, él o ella produce muchas implementaciones y las prueba a todas. La confusión se resuelve si nos damos cuenta de que las partes interesadas externas y el investigador están pensando en diferentes problemas y, por lo tanto, significan algo diferente cuando se habla de una "implementación" (Wieringa, 2014).

Una vez probado el prototipo se lo desplegó en los servidores del grupo de investigación de la Universidad de las Fuerzas Armadas, GrIMPSoftII.

2.3.3.5.Evaluación de la implementación

La validación se contrasta con la evaluación, que es la investigación de un tratamiento según lo aplicado por los interesados en el campo. En el ciclo de ingeniería, la evaluación se realiza después de la implementación. En la evaluación de la implementación, tenemos el beneficio de una visión retrospectiva y podemos utilizar la experiencia de partes interesadas externas con el artefacto implementado para mejorar nuestra teoría de diseño. El objetivo de la investigación de evaluación es investigar cómo interactúan los artefactos implementados con su contexto del mundo real. La investigación de evaluación es una investigación de campo de las propiedades de los artefactos implementados. Los métodos de investigación más utilizados son las encuestas estadísticas y los estudios de casos observacionales, pero la investigación

experimental también se puede usar de otro modo. Un caso especial es la investigación de acción técnica, que es un método para probar un nuevo artefacto en el mundo real usándolo para resolver un problema del mundo real. Esta es una investigación de validación, hecha en el campo. El artefacto todavía está en desarrollo y no es utilizado por las partes interesadas independientemente del contexto de investigación. El artefacto es utilizado por los investigadores para probar sus propiedades en condiciones reales (Wieringa, 2014).

Con el fin de evaluar la utilidad del prototipo se realizó una encuesta, la misma que esta explicada en capítulos posteriores.

2.3.4. Design Science aplicado

En la Figura 4 se muestran las fases y actividades que se aplicaron en la presente investigación.

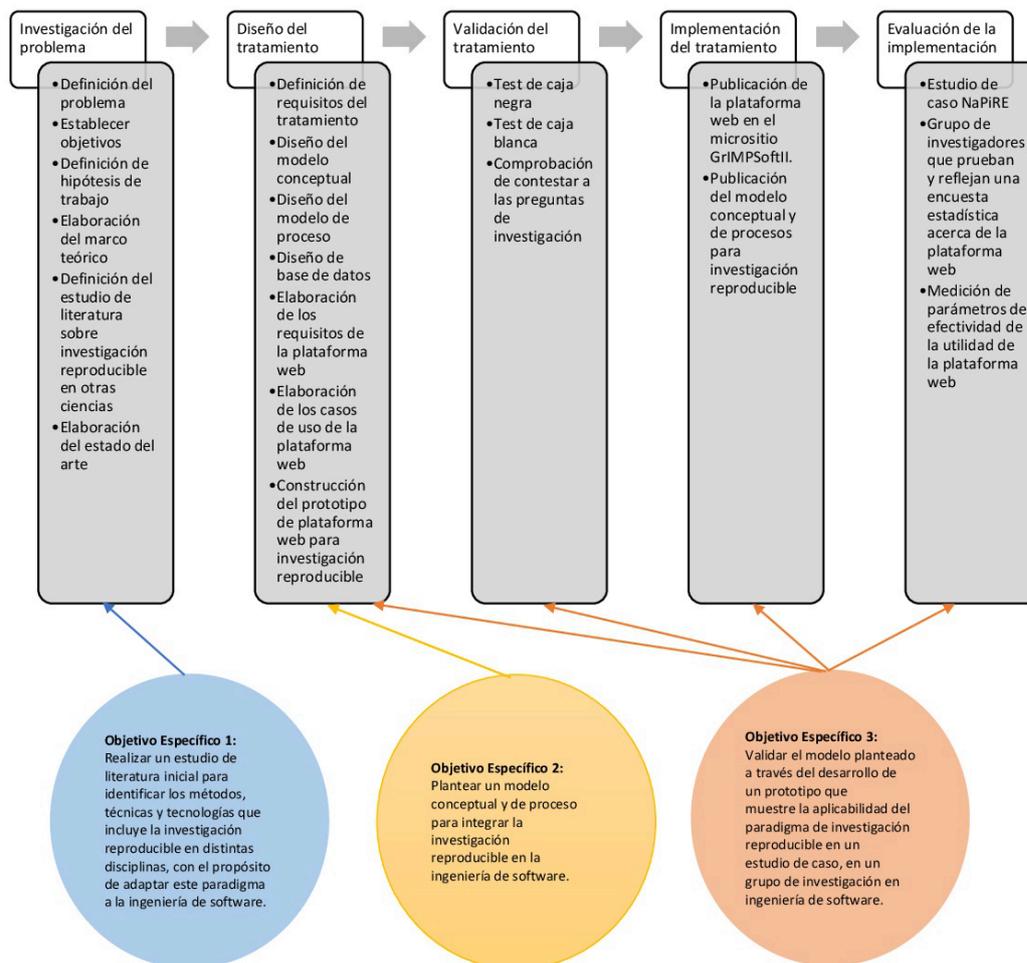


Figura 4. Ciclo Design Science aplicado a la investigación

CAPÍTULO III

MODELOS PROPUESTOS PARA REPRODUCIBILIDAD

Los modelos propuestos son una abstracción del paradigma de investigación reproducible, misma que fue adaptada para ser aplicada a la ingeniería de software. Con los siguientes modelos se pretende explicar la lógica, entidades e interacciones que más adelante conformarán el funcionamiento de este paradigma.

3.1. MODELO CONCEPTUAL DE INVESTIGACIÓN REPRODUCIBLE

El modelo conceptual de investigación reproducible describe los módulos que conforman la herramienta propuesta. En ella se definen procesos, representados por rectángulos, y artefactos o librerías, mismas que se representa con el símbolo de documento. El modelo propuesto en la Figura 5 muestra cinco módulos interconectados entre sí, así como también artefactos externos.

La herramienta propuesta cumple el objetivo de aplicar el paradigma de investigación reproducible a la ingeniería de software, para ello se ha seleccionado a las investigaciones científicas de tipo encuesta, es por ello que como elemento externo se tiene un caso de estudio. Este caso de estudio va a ser el que se ingresa a la herramienta con el fin de que se convierta en un estudio replicable.

El módulo de definición tiene como propósito delimitar y crear la investigación científica de tipo encuesta. Para ello se necesitan cuatro elementos cargados de información. La definición del estudio se convierte en el enfoque que tiene la investigación, en ella se define el problema, los objetivos, la descripción y las preguntas de investigación. Como consecuencia de esta definición se elabora el cuestionario con todas las preguntas que conforman la encuesta del estudio, el cuestionario puede tener varios tipos de preguntas, como son: preguntas abiertas, preguntas cerradas, preguntas de opción única, preguntas de opción múltiple, entre otros. Al aplicar esta encuesta se genera datos y las preguntas abiertas necesitan ser categorizadas antes

de poderse presentar estadísticamente, frente a esta situación se define un “método de tratamiento de datos”; este método consiste en especificar palabras que representan categorías y subcategorías que van a depender del ámbito y carácter de la encuesta y que combinándolos se tiene una etiqueta que categoriza las respuestas abiertas. Finalmente, se proporciona una estructura del reporte de resultados, este documento viene a ser una plantilla donde se plasman los hallazgos encontrados al final de la investigación, la plantilla consiste en varios archivos en formato LaTeX donde mediante el uso de etiquetas personalizadas se indica dónde la herramienta va a colocar los resultados de la encuesta.

El módulo de ejecución consiste básicamente en aplicar la encuesta a la población seleccionada, de modo que se recolecten los datos en crudo. Aquí se aplica la encuesta definida en el módulo anterior. A tiempo que se ejecuta la encuesta, la herramienta se encarga de almacenar los datos para cada estudio en particular.

Con los datos almacenados se pasa al módulo de análisis. Aquí se trata la información, y con el método de tratamiento proveído en el módulo de definición, se categorizan las respuestas a preguntas abiertas, también se definen los parámetros de análisis como filtros de información, entre otros. Todo este análisis es empujado por el servidor de R ya que este ejecuta los scripts pertinentes.

Todos los gráficos y resultados estadísticos generados por el servidor de R son empaquetados por la herramienta y colocados en la estructura del reporte automáticamente.

El módulo final también tiene interacción con elementos externos a la herramienta. A gusto del investigador, este empleará herramientas auxiliares como: programas externos, reportes escritos de otros autores y el compilador de texto LaTeX para dejar integro de información al reporte de resultados. Este reporte, llamado reporte final de resultados, se almacena en el repositorio de investigación reproducible.

Una vez explicado el funcionamiento de los cinco módulos, y comprendiendo la evolución que tiene un caso de estudio dentro de la herramienta de investigación reproducible, en la siguiente sección se explica cómo es el proceso que la aplica en un contexto reproducible.

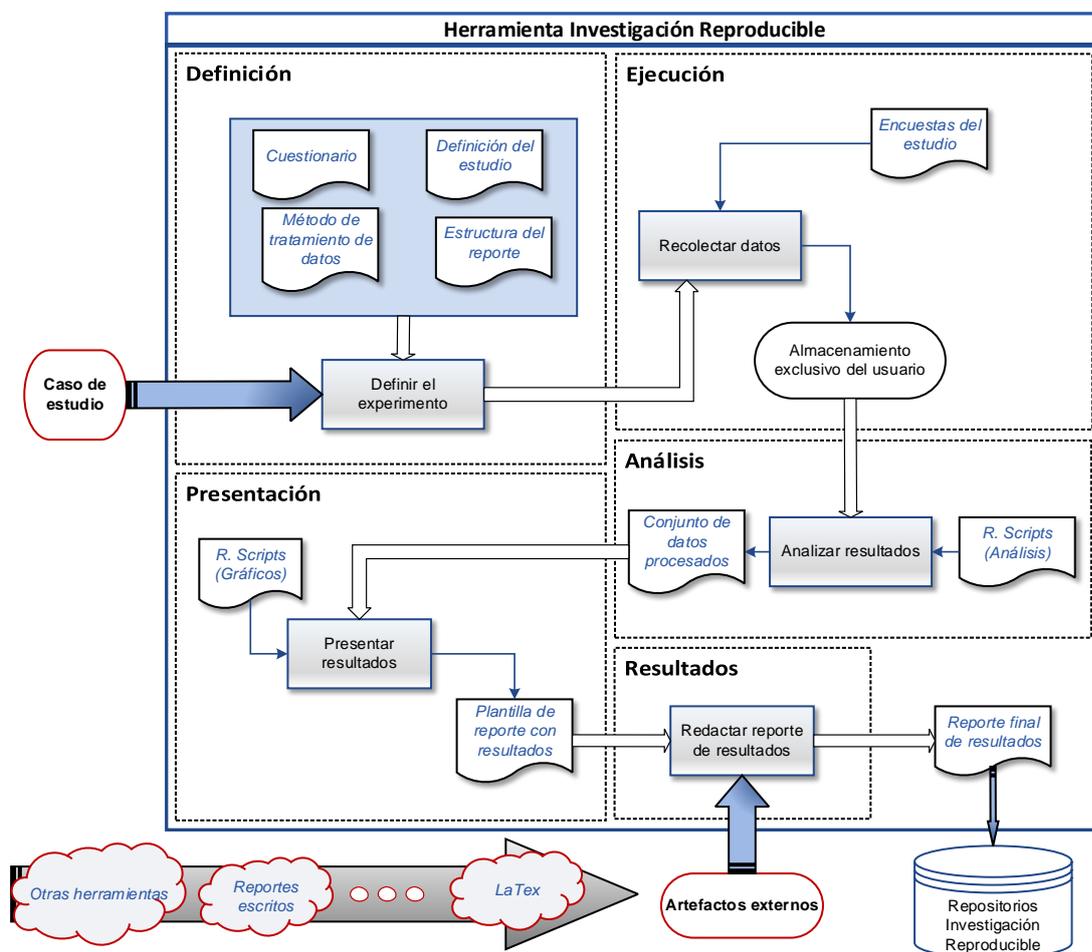


Figura 5. Modelo conceptual de investigación reproducible

3.2. MODELO DE PROCESOS PARA INVESTIGACIÓN REPRODUCIBLE

Para adaptar el paradigma de Investigación Reproducible a la ingeniería de software se elaboró el modelo de procesos que se muestra en la Figura 6. Proceso de Investigación Reproducible

Figura 6, mismo que se utilizó para elaborar el prototipo que permitió aplicar Investigación Reproducible en el caso de estudio NaPiRE.

Este proceso consta de tres subprocessos, que son: Diseñar Estudio, Replicar Estudio y Analizar Conjunto Global de Datos. Todo este proceso tiene algunas fases cíclicas que serán

explicadas a detalle a posterior. De igual modo, los lanes y pools se explican a breves rasgos con el fin de facilitar la explicación de cada subproceso involucrado en el modelo de procesos.

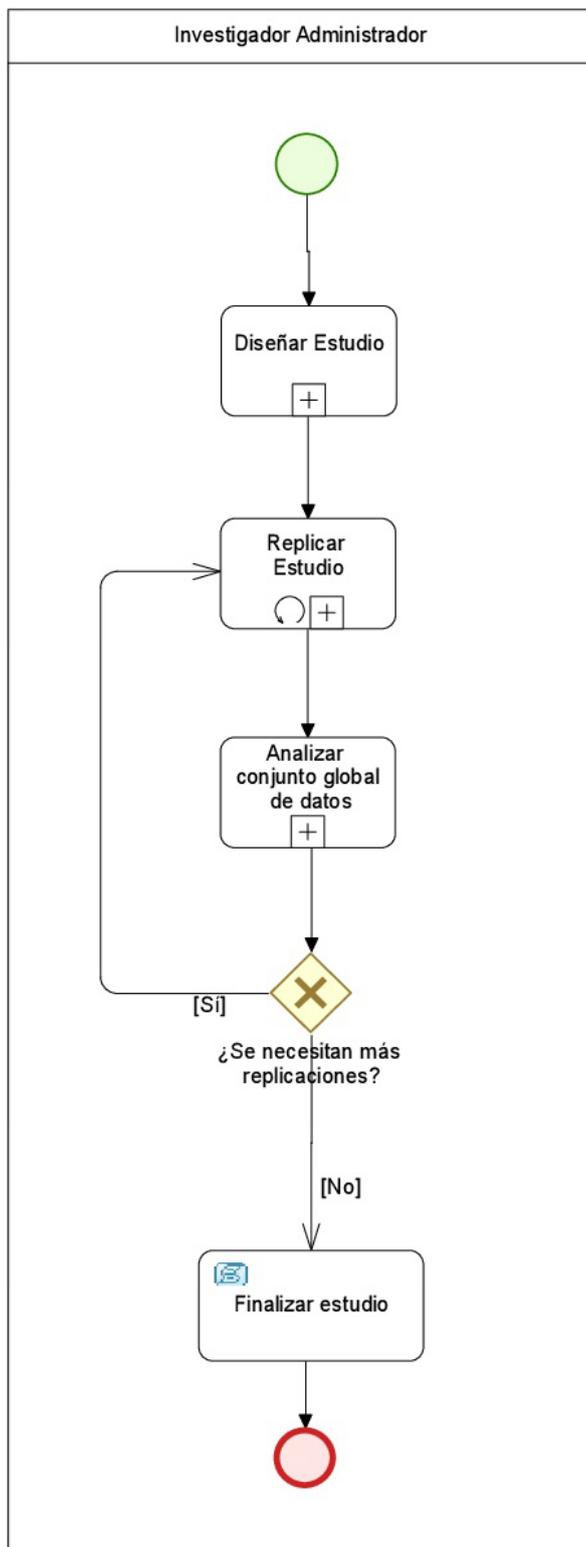


Figura 6. Proceso de Investigación Reproducible

3.2.1. Pools

Se identificaron tres organizaciones, también denominadas pools. Estas organizaciones comprenden el nivel más alto de agrupación de entidades que intervienen en el proceso de Investigación Reproducible, y son:

- Investigación Reproducible: comprende a la organización que se encarga del funcionamiento interno del sistema. En el caso de estudio corresponde a GrIMPSoftII.
- Empresa desarrolladora: comprende a la empresa desarrolladora de software a la que se le aplica la encuesta. En el caso de estudio corresponde a toda empresa a la que se le aplica la encuesta NaPiRE.
- Investigadores externos: comprende a investigadores ajenos al grupo de investigación GrIMPSoftII, que tengan conocimiento del área de investigación en la cual se está desarrollando el estudio que pretende ser publicado y replicado.

3.2.2. Lanes

Dentro de cada organización se diferencia al menos una lane, misma que se conoce como canal. Esta línea representa un participante dentro de la organización, también se lo puede identificar como un rol dentro de todo el proceso.

En el pool de Investigación Reproducible se tiene los siguientes canales:

- Investigador Administrador: corresponde al líder del equipo de investigación, el mismo que tiene los mayores privilegios y responsable de más subprocesos dentro de su organización. Tiene la capacidad de crear nuevos estudios, así como también replicarlos y generar análisis globales de datos en las

replicaciones.

- Investigador Usuario: corresponde a cualquier investigador dentro del equipo de investigación. Es capaz de llevar a cabo replicaciones de estudios y realizar análisis de datos, todo esto en su propio espacio de trabajo.

En el pool de Empresa desarrolladora se tiene el canal denominado “Encuestado”. Este representa a cualquier miembro de la empresa desarrolladora a quien se le aplica la encuesta, puede pertenecer a cualquier departamento organizacional de la empresa, su único requisito es que conozca sobre la temática que la encuesta este abordando. Dentro de todo el proceso de Investigación Reproducible, su razón de ser radica en contestar la encuesta planteada.

El pool Investigador Externo tiene el único canal que se denomina con el mismo nombre. Su razón de ser es realizar retroalimentación al Investigador Administrador acerca del estudio que este está diseñando y publicando.

3.2.3. Subprocesos

3.2.3.1. Diseñar Estudio

Este subproceso comprende la creación de un estudio. Dentro de todo el proceso de Investigación Reproducible se entiende por estudio a un trabajo de investigación cuyo tipo es estrictamente encuesta, también conocido como survey en el idioma inglés. Agrupando las actividades representadas en la Figura 7, se pueden diferenciar cuatro etapas: diseño del estudio, validación del estudio, primera ejecución y publicación del estudio.

La primera etapa, diseño del estudio, busca definir la forma y esencia de la investigación que se pretende replicar, esto comprende: preguntas de investigación, objetivos de la investigación, cuestionario que conforma la encuesta y otros detalles que identifican al estudio.

Posterior a un primer diseño se valida el mismo con ayuda de investigadores externos, quienes analizan todo el planteamiento del estudio y realizan una retroalimentación al

investigador, este a su vez analiza las recomendaciones y las aplica en una mejora al diseño del estudio propuesto. Todas estas actividades se realizan las veces que sea necesario hasta que el investigador esté satisfecho con el diseño de su investigación.

Una fase crucial en el diseño del estudio es la primera ejecución. En esta fase el investigador lleva a cabo toda la investigación de principio a fin, es decir, aplica las encuestas, analiza los datos y saca conclusiones de su primer trabajo. La forma en que el investigador realiza todos los pasos de la investigación darán la pauta para que próximos investigadores que deseen replicar este estudio los vayan siguiendo.

Finalmente, una vez completada toda la investigación por primera vez, el estudio está listo para ser publicado y puesto a disposición a todo investigador interesado en su replicación. Los resultados relevantes que brindó la primera ejecución del estudio también son expuestos a la comunidad científica como muestra latente de hallazgos que generaría una nueva replicación del estudio.

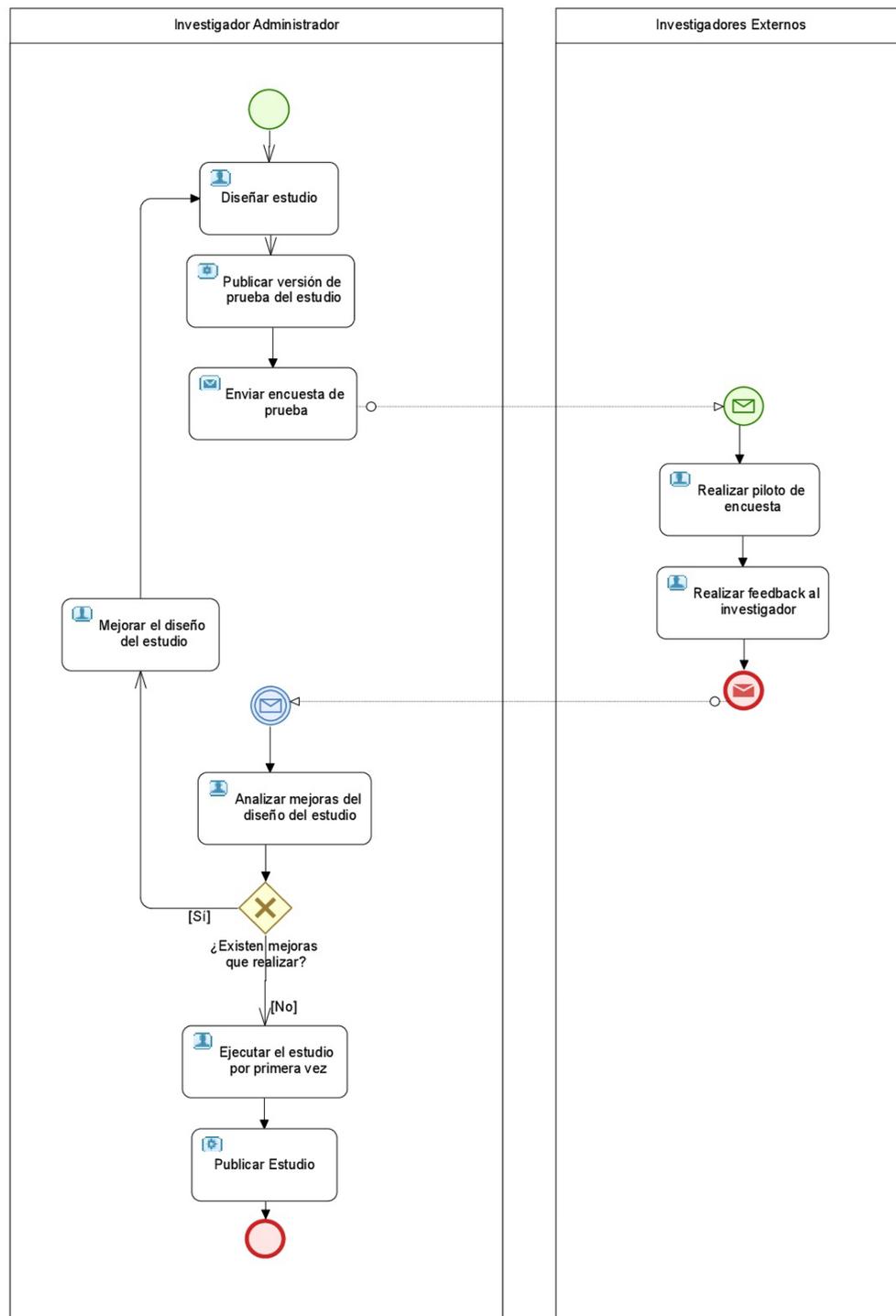


Figura 7. Subproceso Diseñar Estudio

3.2.3.2. Replicar Estudio

Replicar estudio es un subproceso cíclico, el canal que lo inicia es “Investigador Usuario”. Al ser un subproceso cíclico significa que diferentes investigadores van a tener que realizarlo independientemente cada uno. Este subproceso, representado en la Figura 8, está conformado a su vez por dos subprocesos, que son: solicitar autorización (Figura 9) y ejecutar replicación de estudio (Figura 12).

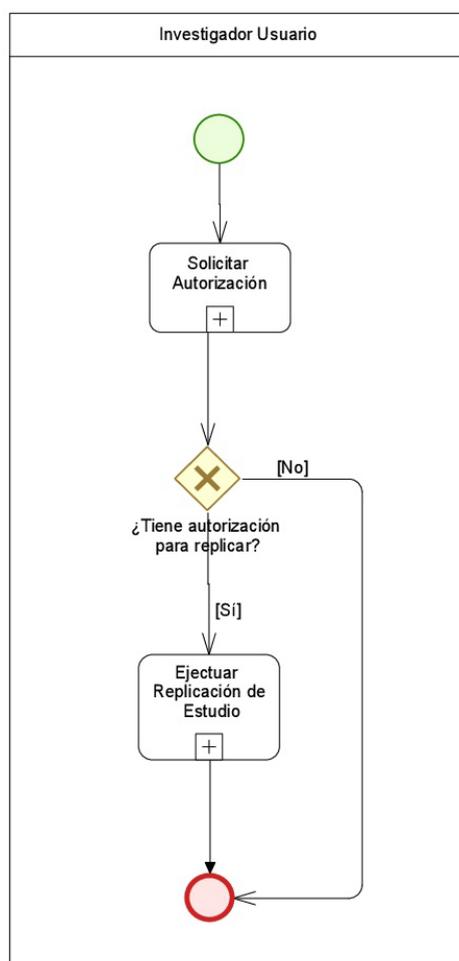


Figura 8. Subproceso Replicar Estudio

El subproceso “Solicitar Autorización”, representado en la Figura 9, conforma la primera etapa del subproceso “Replicar Estudio”. Su finalidad es otorgar dos permisos. El primer permiso corresponde al acceso a la plataforma web mediante la creación de un usuario,

específicamente el usuario tipo “investigador usuario”, el flujo del proceso se lo puede observar en la Figura 10. El segundo permiso que otorga este subproceso es la autorización para replicar un estudio, es decir, una vez que el usuario tiene un usuario para ingresar a la plataforma web, este puede seleccionar un estudio que se encuentre disponible para realizar su propia replicación del mismo; las actividades involucradas en este subproceso se muestran en la Figura 11. Los subprocesos de autorización funcionan con un comportamiento similar, ambos son aprobados por el canal “Investigador Administrador” mediante el análisis de la información proporcionada por el canal “Investigador Usuario”, en ella se evalúa si la información proporcionada cumple con los requerimientos necesarios para que el candidato que solicita los permisos sea apto para recibir dichos privilegios, en caso de no ser suficiente se procede a solicitar más información o, en el peor de los casos, denegar los permisos, en cuyo caso se repite el subproceso “Replicar Estudio” con un nuevo participante. La aprobación o denegación de los permisos se informa a través del correo electrónico del “Investigador Usuario” que los solicita.

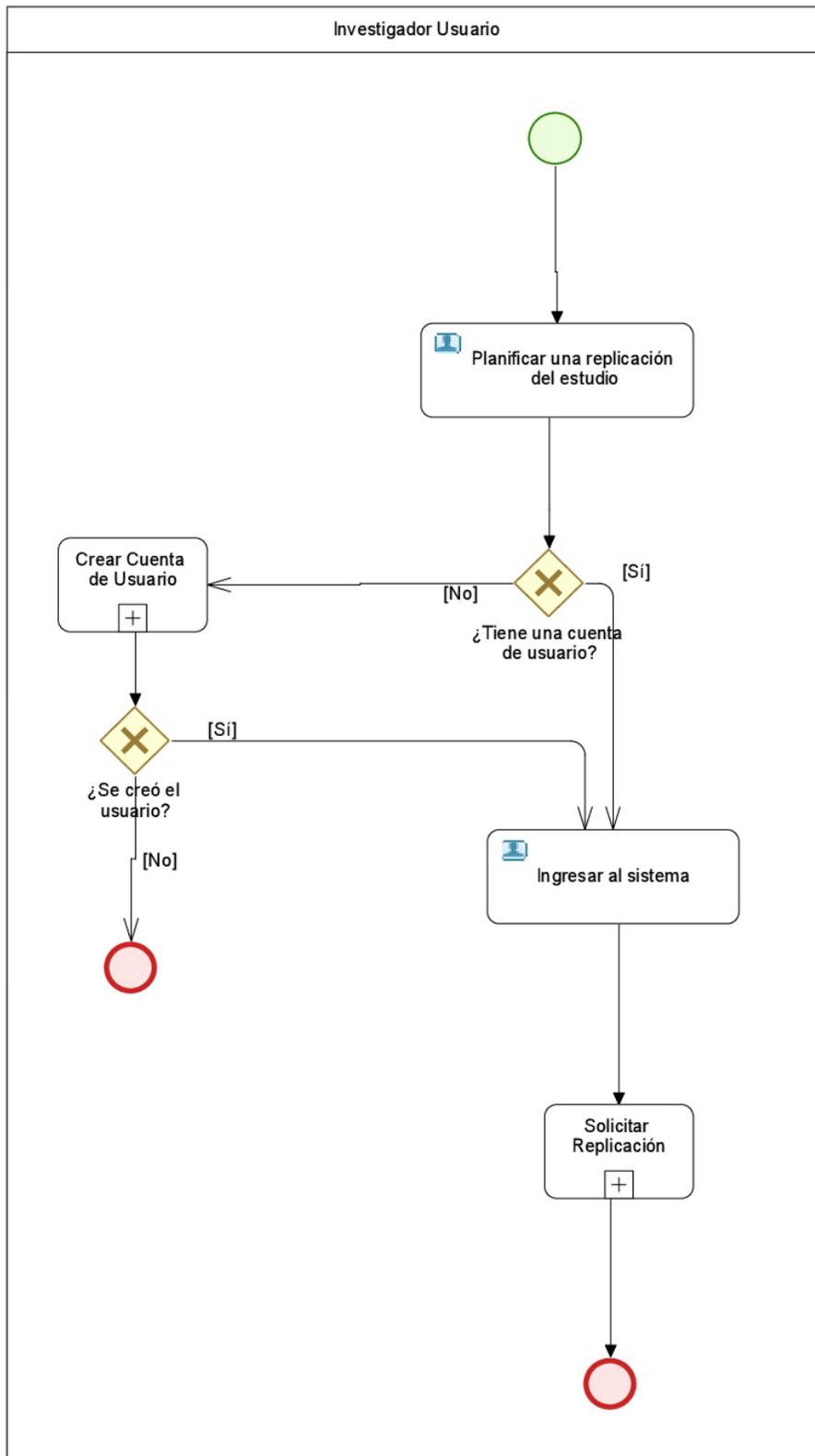


Figura 9. Subproceso Solicitar Autorización

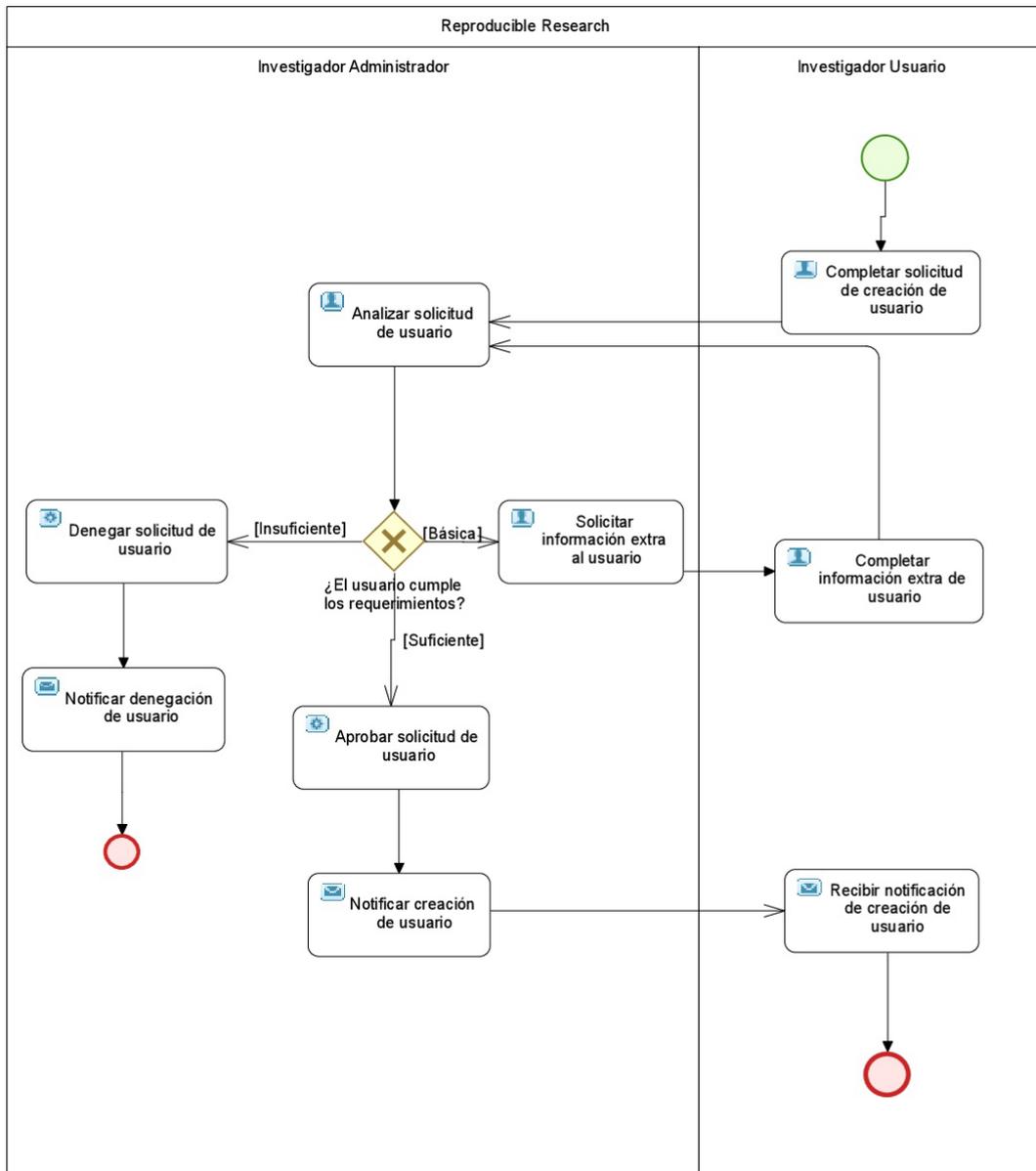


Figura 10. Subproceso Crear Cuenta de Usuario

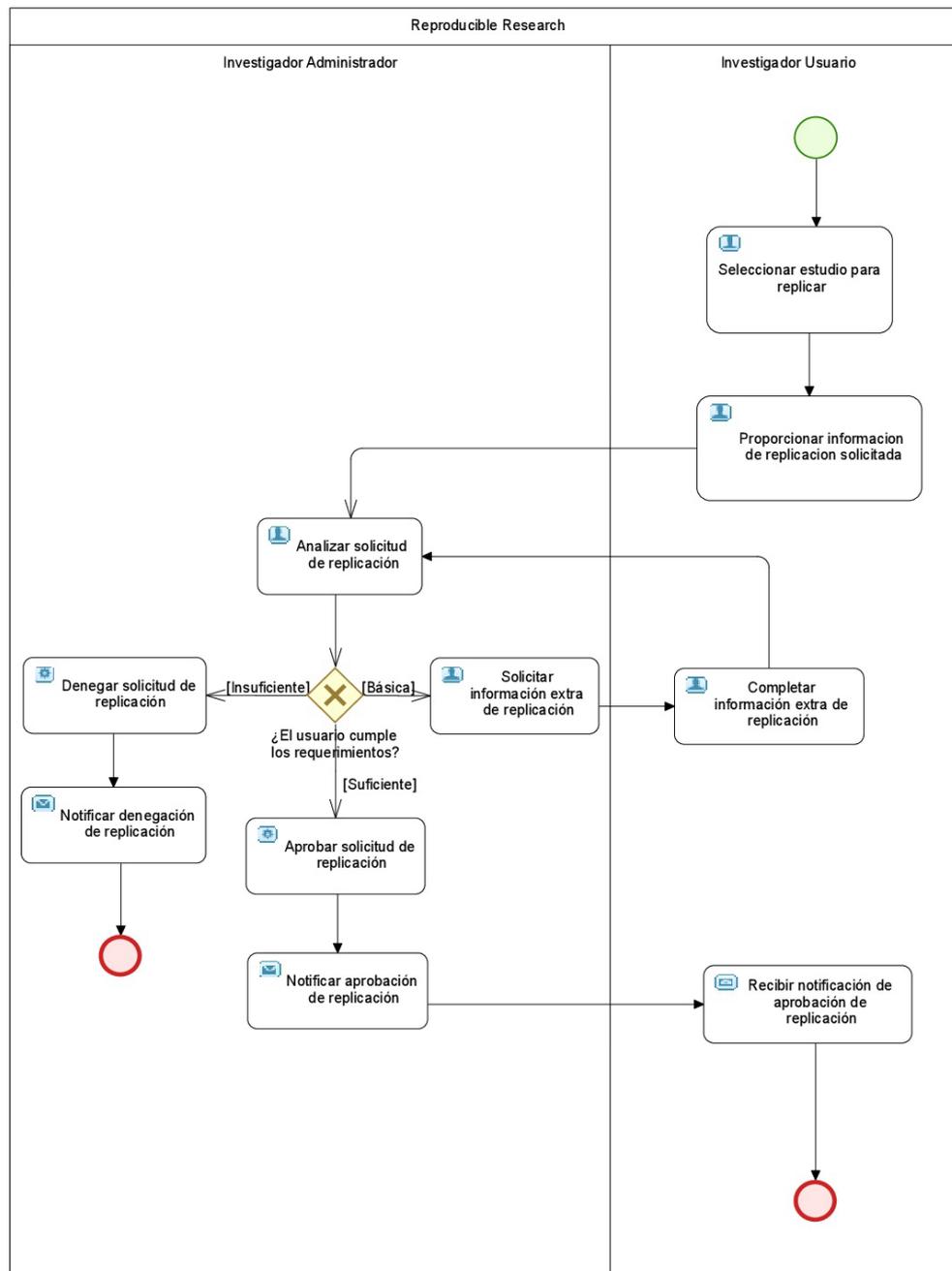


Figura 11. Subproceso Solicitar Replicación

Continuando con el flujo del subproceso “Replicar Estudio” (véase Figura 8), y una vez que se han conseguido los permisos necesarios para replicar un estudio viene un subproceso vital. Este subproceso es muy importante, pero no el que le dará total sentido a la investigación reproducible. El subproceso “Ejecutar Replicación del Estudio” consiste en ejecutar la investigación, es decir, aplicar la encuesta a nuevas empresas desarrolladoras que no fueron

usadas en la ejecución inicial realizada por el canal “Investigador Administrador”. Este subproceso tiene actividades más elaboradas, por lo que se va a desglosar su utilidad.

En la actividad “Diseñar la invitación (Flyer)” se diseña el volante, flyer o invitación que se utiliza para invitar a las empresas desarrolladoras a que contesten a la encuesta. Aquí el investigador usuario puede personalizar ciertos datos informativos que son propios de su instancia de replicación.

Con ayuda de la plataforma web, en la actividad “Difundir el estudio”, se ingresan los correos de los encuestados, mismos a los que se envía un email con la invitación a participar en la encuesta adjuntando el link donde está alojada la misma. La plataforma web se encarga de almacenar las respuestas a la encuesta y el investigador que está replicando el estudio puede ingresar más correos de los encuestados, de ser necesario recolectar más información.

Una vez recolectada toda la información necesaria, viene la fase de análisis, de esto trata la actividad “Definir requerimientos y parámetros de análisis”. Previo a definir cómo se quiere mostrar la información, es necesario categorizar las respuestas en preguntas abiertas, esto se lo logra mediante categorías que el investigador administrador definió al momento de diseñar su estudio, es así que una a una se clasifican las respuestas abiertas de forma que puedan ser presentadas como resultados tabulados. Una vez categorizadas las respuestas a preguntas abiertas, el investigador tiene la opción de aplicar filtros a la información recolectada, así como también definir la forma en que se presentarán las gráficas de los resultados del cuestionario de la encuesta.

Con ayuda del servidor R y LaTeX, se analiza, procesa y presenta los resultados al cuestionario de la encuesta. La plataforma web entrega un archivo comprimido con muchos archivos a forma de plantilla de un reporte de resultados. Esta plantilla está en formato LaTeX con algunas secciones a manera de paper, aquí el investigador usuario irá trabajando localmente

en su computadora redactando el reporte de resultados que ha brindado su replicación del estudio tipo encuesta que ha replicado. Este reporte refleja todos los hallazgos que ha tenido el investigador, así como también conclusiones y recomendaciones. Una vez que el investigador haya completado su reporte debe subirlo tal cual la plataforma le entregó el comprimido, esta actividad se la realiza ya sea por la plataforma web o con ayuda de un ejecutable “jar” que viene incluido en el archivo de resultados.

Llegado hasta este punto, el investigador tiene opción de modificar los parámetros de análisis, en caso de que haya decidido dar otro enfoque a la información obtenida en su replicación, o en caso de estar satisfecho, dar por concluida su replicación.

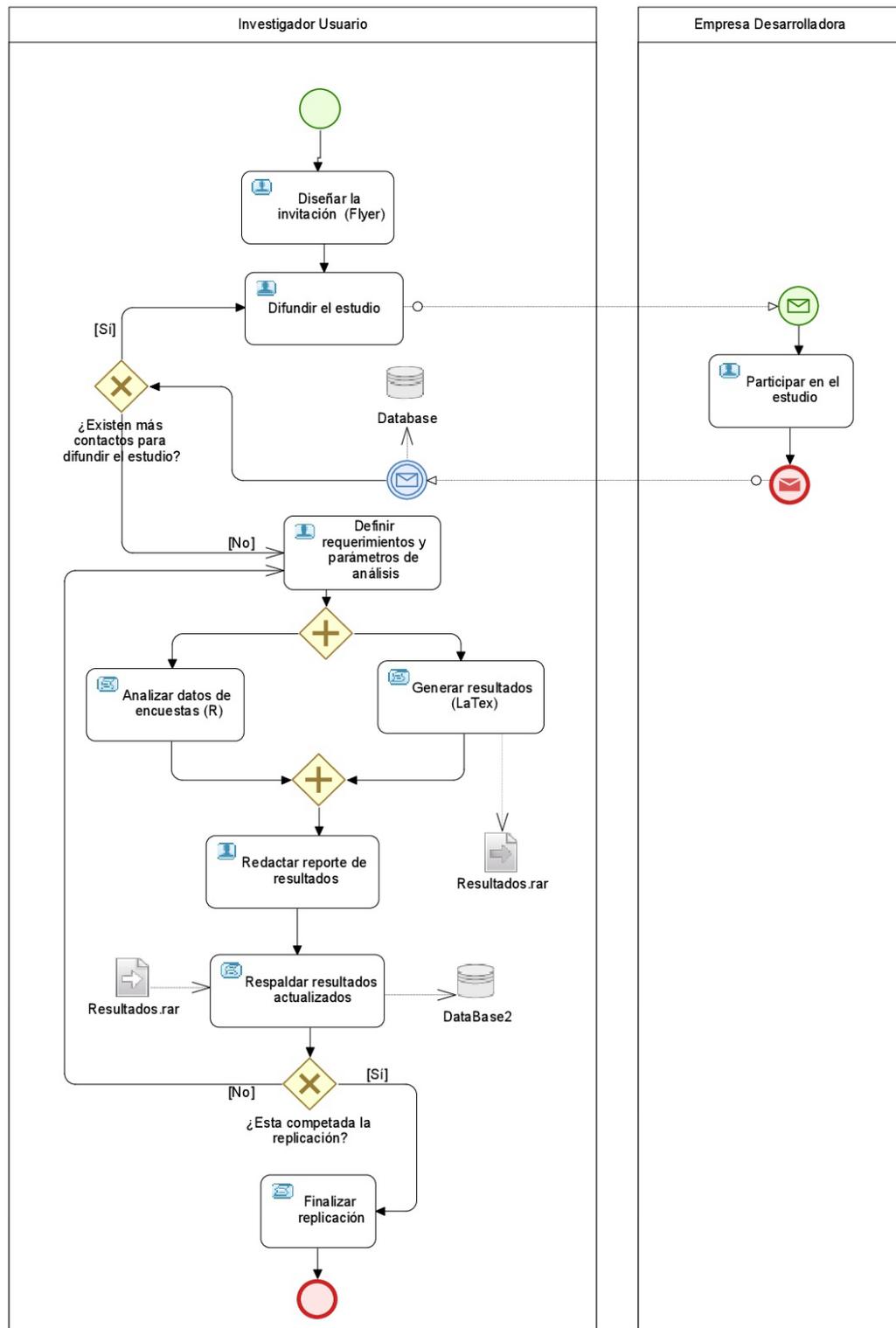


Figura 12. Subproceso Ejecutar Replicación de Estudio

3.2.3.3. Analizar Conjunto Global de Datos

El subproceso “Analizar Conjunto Global de Datos” corresponde a la última pieza fundamental dentro del proceso de Investigación Reproducible. La Figura 13 ilustra las actividades involucradas, en ella se observa cierta similitud con el subproceso “Ejecutar Replicación de Estudio” (ver Figura 12) debido a que involucra análisis de datos, seguido del desarrollo del reporte de resultados. La primera actividad se caracteriza por la selección de los datos de las replicaciones del estudio, dando opción al investigador de elegir qué datos desea procesar, además se utiliza el principio de Pareto para analizar la información.

Este subproceso es muy importante ya que permite condensar los hallazgos de todas las instancias de replicación del estudio en cuestión, de manera que se puede aumentar la confianza en los resultados.

Finalmente, continuando con el proceso de investigación reproducible (ver Figura 6) se tiene una puerta de enlace en la que el investigador administrador decidirá si los datos obtenidos de todas las replicaciones son suficientes o es necesario aguardar por más investigadores que vayan a replicar el estudio. En caso de que se desee realizar más replicaciones el proceso continúa con el subproceso “Replicar Estudio” (ver Figura 8), caso contrario se da por concluido el estudio y no se puede realizar más replicaciones del mismo.

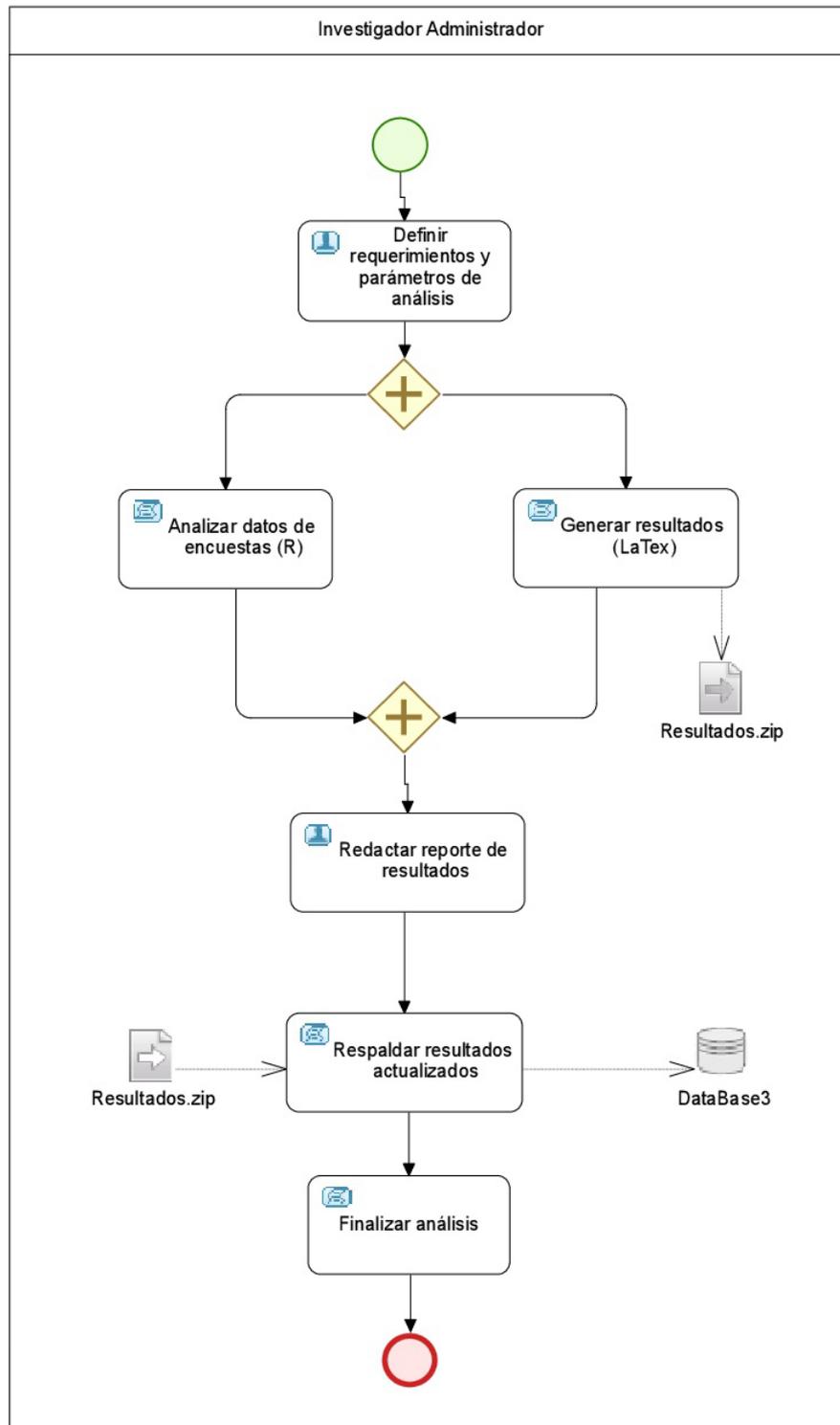


Figura 13. Subproceso Analizar Conjunto Global de Datos

CAPÍTULO IV

DISEÑO DEL PROTOTIPO DE PLATAFORMA WEB

En este capítulo se exponen los diseños que fueron utilizados para el desarrollo del prototipo de Investigación Reproducible. Entre los diseños utilizados se tiene: diseño de base de datos, especificación de requisitos y caso de estudio, arquitectura de software, SCRUM e interfaces de usuario.

4.1 DISEÑO DE BASE DE DATOS

El almacenamiento de datos es arista fundamental en el desarrollo de la plataforma web, por lo que se utilizaron los tres modelos: modelo conceptual, modelo lógico y modelo físico.

4.1.1. Modelo conceptual

El diagrama de la Figura 14 expresa las entidades y relaciones que interactúan en la base de datos vistos de manera conceptual, es decir, no a un nivel más claro de abstracción.

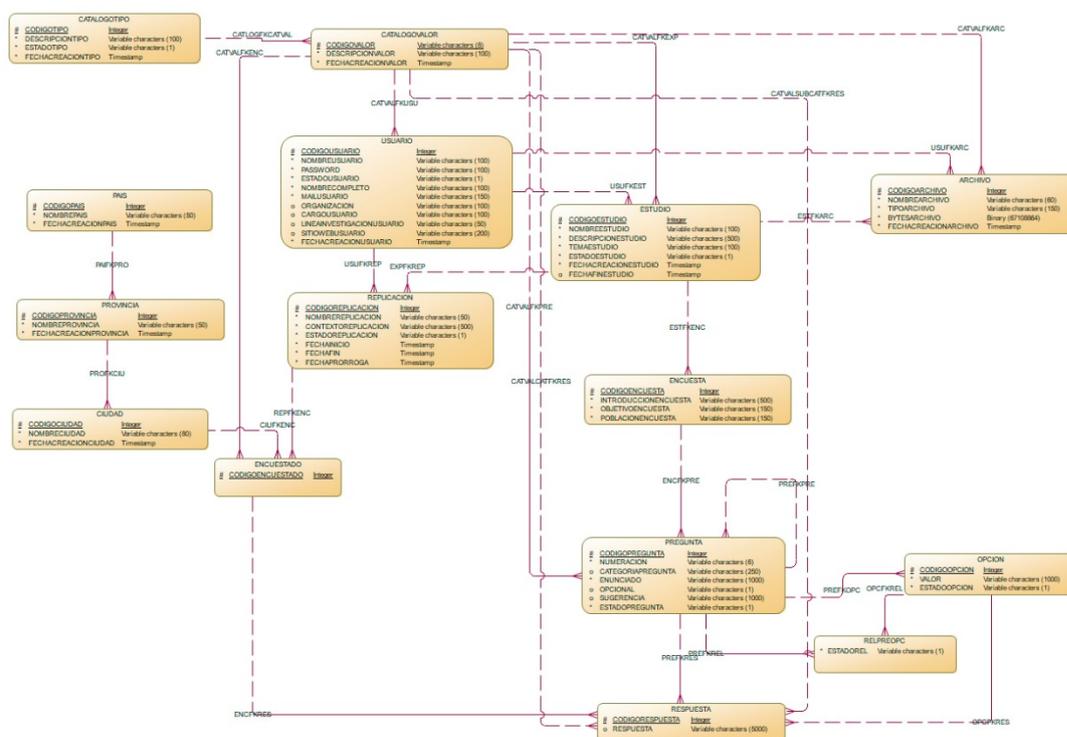


Figura 14. Modelo conceptual de base de datos

4.2.1 Historias de Usuario

Tabla 4

Historia de usuario 1

| | | |
|-------------------------------|-------------|--|
| Id. usuario | Historia de | TFR-001 |
| Prioridad | | Alta |
| Nombre | | Login del sistema |
| Rol | | Como usuario del sistema |
| Característica /Funcionalidad | | Necesito tener una cuenta e ingresar a la plataforma, con la finalidad de utilizar las funcionalidades de la plataforma web |
| Validación | | <ol style="list-style-type: none"> 1. Usuario o contraseña incorrectos: En caso que un usuario intente ingresar a la plataforma, cuando su usuario o contraseña sean incorrectos, el sistema deberá informar al usuario de su error y pedirle que vuelva a ingresar. 2. Usuario y contraseña correcto: En caso que un usuario intente ingresar a la plataforma, cuando su usuario y contraseña sean correctos, el sistema debe permitir el acceso a la plataforma y mostrar el menú principal con las opciones disponibles según el rol del usuario. |

Tabla 5

Historia de usuario 2

| | | |
|-------------------------------|-------------|--|
| Id. usuario | Historia de | TFR-002 |
| Prioridad | | Alta |
| Nombre | | Gestión de estudios (investigaciones) |
| Rol | | Como un investigador administrador |
| Característica /Funcionalidad | | Necesito gestionar los estudios, con la finalidad de añadir y manipular los estudios que serán replicados. |
| Validación | | <ol style="list-style-type: none"> 1. Crear nuevo estudio: En caso que un usuario cree un nuevo estudio, cuando llene los campos necesarios correctamente, el sistema debe almacenar dicho estudio y ponerlo a disposición de los investigadores que quieran replicarlo 2. Eliminar un estudio: En caso que un usuario elimine un estudio, cuando seleccione el estudio que desea eliminar el sistema debe eliminarlo del catálogo disponible de estudios replicables. 3. Modificar un estudio: En caso que un usuario modifique un estudio, cuando cambie algún parámetro del estudio el sistema debe modificar sus valores y presentar la última versión en la plataforma. |

Tabla 6

Historia de usuario 3

| | | |
|-------------|-------------|--|
| Id. usuario | Historia de | TFR-003 |
| Prioridad | | Alta |
| Nombre | | Gestión de cuestionarios de los estudios |

CONTINÚA 

| | |
|-------------------------------|---|
| Rol | Como un investigador administrador |
| Característica /Funcionalidad | Necesito gestionar los cuestionarios de las investigaciones, con la finalidad de crear encuestas que pertenezcan a un estudio específico |
| Validación | <ol style="list-style-type: none"> 1. Añadir preguntas: En caso que un usuario añada preguntas a un estudio, cuando quiera crear un cuestionario para el estudio específico, el sistema debe permitir crear cuantas preguntas sea necesario. 2. Eliminar preguntas: En caso que un usuario elimine preguntas de un estudio, cuando quiera quitar preguntas del estudio, el sistema debe permitirle eliminar las preguntas que desee. 3. Modificar preguntas: En caso que un usuario modifique el cuestionario, cuando cambie los parámetros de una pregunta, el sistema debe permitirle cambiar siempre que tenga coherencia con el tipo de pregunta. |

Tabla 7
Historia de usuario 4

| | |
|-------------------------------|--|
| Id. Historia de usuario | TFR-004 |
| Prioridad | Alta |
| Nombre | Solicitar permiso de replicación |
| Rol | Como investigador usuario |
| Característica /Funcionalidad | Necesito solicitar el permiso y los recursos para replicar un estudio, con la finalidad de hacer una replicación nueva de dicho estudio. |
| Validación | <ol style="list-style-type: none"> 1. Pedir permiso de replicación: En caso que el usuario seleccione un estudio disponible para su replicación, cuando haya seleccionado y aceptado la replicación, el sistema debe informar al investigador administrador y éste deberá aprobar o rechazar la solicitud. |

Tabla 8
Historia de usuario 5

| | |
|-------------------------------|---|
| Id. Historia de usuario | TFR-005 |
| Prioridad | Media |
| Nombre | Difundir link de encuesta |
| Rol | Como investigador usuario |
| Característica /Funcionalidad | Necesito enviar la encuesta a todas las personas que quiero encuestar, con la finalidad de recolectar información que me permita replicar el estudio |
| Validación | <ol style="list-style-type: none"> 1. Enviar link: En caso que el usuario quiera enviar el link de las encuestas, cuando haya ingresado los correos de las personas a las que quiere enviar la encuesta, el sistema debe enviarles un mail con el link para que puedan responder el cuestionario. |

Tabla 9
Historia de usuario 6

| | |
|-------------------------------|---|
| Id. Historia de usuario | TFR-006 |
| Prioridad | Alta |
| Nombre | Clasificar respuestas a preguntas abiertas |
| Rol | Como un investigador usuario |
| Característica /Funcionalidad | Necesito categorizar las respuestas a preguntas abiertas, con la finalidad de tabular la información y poder analizarla. <ol style="list-style-type: none"> Categorizar preguntas: En caso que el usuario vaya a categorizar preguntas abiertas, cuando ya quiera ver los resultados de la encuesta, el sistema debe darle opciones de acuerdo al cuestionario, las mismas que serán elegidas por el usuario y para que organice las preguntas abiertas en las categorías expuestas |
| Validación | |

Tabla 10
Historia de usuario 7

| | |
|-------------------------------|--|
| Id. Historia de usuario | TFR-007 |
| Prioridad | Alta |
| Nombre | Análisis de encuestas |
| Rol | Como un investigador usuario o investigador administrador |
| Característica /Funcionalidad | Necesito que las preguntas de la encuesta sean analizadas estadísticamente, con la finalidad de elaborar un informe técnico de la investigación. <ol style="list-style-type: none"> Análisis simple: En caso que el usuario quiera ver los resultados de la encuesta, cuando entre al apartado de análisis, el sistema debe mostrarle los resultados de su replicación y permitirle aplicar filtros a la información analizada. Análisis compuesto: En caso que el administrador quiera ver los resultados de varias encuestas, cuando haya seleccionado varias replicaciones de su estudio, el sistema debe mostrarle los resultados de su replicación y permitirle aplicar filtros a la información analizada. |
| Validación | |

Tabla 11
Historia de usuario 8

| | |
|-------------------------------|--|
| Id. Historia de usuario | TFR-008 |
| Prioridad | Alta |
| Nombre | Respaldo de informe técnico en la nube |
| Rol | Como un investigador usuario |
| Característica /Funcionalidad | Necesito respaldar los resultados de la investigación en la nube, con la finalidad de tener actualizados los avances del informe técnico <ol style="list-style-type: none"> Trabajo en computadora local: En caso que el usuario esté trabajando el informe técnico en su computadora, cuando ya tenga los |
| Validación | |

resultados de la encuesta aplicada, el sistema le permite subir sus documentos a la plataforma y descargarlos cuando los necesite.

2. **Descargar archivos de investigación:** En caso que el usuario ya haya subido una versión de sus archivos trabajados, cuando ya haya subido una versión antigua, el sistema le permite bajarse sus archivos para continuar trabajándolos.
3. **Subir más versiones de archivos de investigación:** En caso que el usuario ya haya subido una versión de sus archivos trabajados, cuando quiera subir una nueva versión, el sistema le permite subir sus documentos y sobrescribir la última versión subida

Tabla 12
Historia de usuario 9

| | | |
|-------------------------------|-------------|---|
| Id. usuario | Historia de | TFR-009 |
| Prioridad | | Alta |
| Nombre | | Creación de usuarios |
| Rol | | Como un usuario de la plataforma |
| Característica /Funcionalidad | | Necesito crear un usuario, con la finalidad de acceder a las funcionalidades de la plataforma web |
| Validación | | <ol style="list-style-type: none"> 1. Solicitar creación de usuario: En caso que el usuario entre al login de la plataforma, cuando seleccione crear nuevo usuario, el sistema debe solicitarle los datos necesarios e informar al investigador administrador |

Tabla 13
Historia de usuario 10

| | | |
|-------------------------------|-------------|---|
| Id. usuario | Historia de | TFR-010 |
| Prioridad | | Alta |
| Nombre | | Gestión de usuarios |
| Rol | | Como investigador administrador |
| Característica /Funcionalidad | | Necesito gestionar usuarios, con la finalidad de otorgar o denegar el acceso a la plataforma |
| Validación | | <ol style="list-style-type: none"> 1. Aceptar usuario: En caso que un usuario haya solicitado una cuenta en la plataforma, cuando haya sido enviado el mail de aprobación al administrador, el sistema debe aprobar/denegar la creación de usuario, según haya decidido el administrador. 2. Eliminar usuario: En caso que el administrador quiera eliminar un usuario, cuando haya entrado a la gestión de usuarios, el sistema debe inhabilitar al usuario seleccionado por el administrador. |

Tabla 14
Historia de usuario 11

| | |
|-------------------------------|---|
| Id. Historia de usuario | TFR-011 |
| Prioridad | Alta |
| Nombre | Gestión de replications |
| Rol | Como investigador administrador |
| Característica /Funcionalidad | Necesito gestionar las replications, con la finalidad de otorgar o denegar replications |
| Validación | <ol style="list-style-type: none"> Aceptar replicación: En caso que un usuario haya solicitado una replicación de un estudio, cuando el administrador ha recibido el mail de aprobación, el sistema debe otorgar o denegar el permiso de replicación según haya decidido el administrador. Terminar replicación: En caso que el administrador desee acabar una replicación de un usuario, cuando vaya al apartado de gestión de replications, el sistema debe terminar una replicación y ponerla como inactiva. |

Tabla 15
Historia de usuario 12

| | |
|-------------------------------|---|
| Id. Historia de usuario | TFR-012 |
| Prioridad | Alta |
| Nombre | Análisis de encuestas |
| Rol | Como un investigador administrador |
| Característica /Funcionalidad | Necesito publicar los cuestionarios en modo de prueba, con la finalidad de corregir errores en el diseño de la encuesta |
| Validación | <ol style="list-style-type: none"> Estudio en modo prueba: En caso que el investigador administrador quiera poner en modo de prueba su cuestionario, cuando haya creado un nuevo estudio y tenga preguntas su cuestionario, el sistema debe publicar el cuestionario en modo prueba. El link se envía por mail a los investigadores que probarán el cuestionario. Publicar estudio definitivo: En caso que el investigador administrador ya quiera publicar el estudio, cuando ya haya probado su cuestionario en modo prueba, el sistema arrancará el estudio y lo publicará para que sea replicado por otros investigadores usuarios. |

Tabla 16
Historia de usuario 13

| | |
|-------------------------------|---|
| Id. Historia de usuario | TFR-013 |
| Prioridad | Alta |
| Nombre | Análisis conjunto de varias replications |
| Rol | Como un investigador administrador |
| Característica /Funcionalidad | Necesito analizar el conjunto de replications, con la finalidad de generar conocimiento del conjunto de datos |

| | |
|------------|---|
| Validación | <ol style="list-style-type: none"> 1. Agrupar 2 o más estudios: En caso que el investigador administrador quiera analizar dos resultados de replicaciones conjuntamente, cuando haya seleccionado dos o más estudios, el sistema mezclará las dos fuentes de datos para analizarlos como uno solo. 2. Generar reporte conjunto: En caso que el investigador administrador quiera generar un reporte conjunto, cuando ya haya analizado dos o más replicaciones, el sistema le entregará la plantilla para completar el informe técnico así como también le permitirá subirlo a la plataforma web. |
|------------|---|

4.3 REQUISITOS DEL PROTOTIPO

4.3.1 Especificación de requisitos

4.3.1.1. Requisitos funcionales

En la plataforma web desarrollada se identifican los siguientes actores que interactúan con el sistema:

- Investigador administrador: persona líder dentro del grupo de investigación. Tiene las facultades de: gestionar estudios, otorgar permisos y replicar estudios.
- Investigador usuario: persona que pertenece al grupo de investigación. Tiene las facultades de replicar un estudio existente en el sistema.
- Encuestado: persona que pertenece a una empresa desarrolladora de software. Tiene la facultad de contestar a las encuestas pertenecientes a la replicación de un estudio específico.

Los requisitos funcionales de la plataforma web se encuentran detallados a continuación.

Requisito funcional 1

El requisito funcional RR-RF001, denominado “Gestionar un estudio”, se encuentra detallado en la Tabla 17.

Tabla 17
Requisito funcional 1

| | |
|---------------------|---|
| Id. Requerimiento | RR-RF001 |
| Nombre | Gestionar un estudio |
| Actor | Investigador Administrador. |
| Descripción | La aplicación web debe permitir agregar los datos necesarios para la gestión de un estudio. El estudio debe ser específicamente de tipo encuesta. Adicionalmente el estudio requiere un diseño de la investigación que deberá ser proporcionado a la aplicación, el mismo que tendrá un periodo de prueba en el que podrá ser modificado y probado por otros investigadores. Formulario con los campos a llenar: |
| Entradas | <ul style="list-style-type: none"> – Nombre del estudio – Fecha de creación – Descripción del estudio. – Estado del estudio |
| Salidas | <p>-La información será almacenada en la tabla <<ESTUDIO>> de la base de datos.</p> <p>-Se mostrará una ventana con un mensaje de confirmación de la acción realizada. Si el usuario confirma la acción se procede a ejecutarla.</p> <p>1) Seleccionar el apartado de “Estudio”.</p> |
| Proceso | 2) Llenar el formulario que se desplegará. |
| Precondiciones | 3) Seleccionar la opción de envío de datos. Haber ingresado a la aplicación web con el rol de Investigador Administrador. El estudio será guardado, modificado o eliminado con éxito en la base de datos. |
| Post condiciones | |
| Efectos Colaterales | Permite agregar preguntas a la encuesta del Estudio creado. Si el usuario no ingresa todos los campos obligatorios o llena incorrectamente, el sistema no permitirá que el estudio sea almacenado o eliminado. |
| Prioridad | Media |

Requisito funcional 2

El requisito funcional RR-RF002, denominado “Gestionar Cuestionario”, se encuentra detallado en la Tabla 18.

Tabla 18
Requisito funcional 2

| | |
|---------------------|--|
| Id. Requerimiento | RR-RF002 |
| Nombre | Gestionar Cuestionario |
| Actor | Investigador Administrador |
| Descripción | La aplicación web debe permitir al usuario agregar, modificar o eliminar preguntas que conforman la encuesta, mismas que pertenecen a un estudio específico. Formulario de ingreso con los campos a llenar: |
| Entradas | <ul style="list-style-type: none"> - Enunciado de la pregunta - Opciones de respuesta |
| Salidas | <p>- Las preguntas serán almacenadas en la tabla <<PREGUNTA>>, mientras que cada posible opción a la respuesta será almacenada en la tabla <<OPCIONES>>, de la base de datos.</p> <p>-Se mostrará una ventana con el mensaje de confirmación. 1) Seleccionar un Estudio.</p> |
| Proceso | <p>2) Llenar el formulario que se desplegará para gestionar preguntas.</p> <p>3) Seleccionar la opción de enviar información. Haber ingresado a la aplicación web con el rol de Investigador Administrador.</p> |
| Precondiciones | Tener un Estudio creado. |
| Post condiciones | La solución al problema será guardado con éxito en la base de datos. |
| Efectos Colaterales | Si el usuario no ingresa todos los campos obligatorios o llena incorrectamente, el sistema no permitirá que el cuestionario sea gestionado. |
| Prioridad | Media |

Requisito funcional 3

El requisito funcional RR-RF003, denominado “Solicitar Replicación de Estudio”, se encuentra detallado en la Tabla 19.

Tabla 19
Requisito funcional 3

| | |
|-------------------|---|
| Id. Requerimiento | RR-RF003 |
| Nombre | Solicitar Replicación de Estudio |
| Actor | Investigador Usuario, Investigador Administrador |
| Descripción | La aplicación web debe permitir al usuario apartar su propio entorno basado en un Estudio ya creado para poder ejecutar sus propias tareas. |
| Entradas | <ul style="list-style-type: none"> - Estudio a ser replicado. - Fecha de creación. |
| Salidas | - Registro de replicación de estudio que se almacena en la tabla <<REPLICACION>> de la base de datos. |

CONTINÚA 

| | |
|---------------------|---|
| Proceso | 1) Seleccionar el apartado de “Replicaciones”. 2) Aceptar condiciones de replicación. 3) Enviar confirmación. |
| Precondiciones | Debe existir y estar activo al menos un Estudio. |
| Post condiciones | La aprobación queda pendiente en manos del administrador de la aplicación web, específicamente en el requisito RR-RF0011. |
| Efectos Colaterales | |
| Prioridad | Alta. |

Requisito funcional 4

El requisito funcional RR-RF004, denominado “Ejecutar Encuesta”, se encuentra detallado en la Tabla 20.

Tabla 20
Requisito funcional 4

| | |
|---------------------|---|
| Id. Requerimiento | RR-RF004 |
| Nombre | Ejecutar Encuesta |
| Actor | Encuestado. |
| Descripción | La aplicación web debe permitir al usuario ejecutar la encuesta para recolectar los datos para el estudio. La aplicación debe permitir al usuario enviar el link donde los encuestados puedan llenar la información requerida. Datos del encuestado. |
| Entradas | Preguntas de la encuesta. |
| Salidas | -Las respuestas a la encuesta son almacenados en la tabla <<RESPUESTA>>, en la base de datos. 1) Abrir el link brindado por el Investigador |
| Proceso | 2) Contestar cada pregunta de la encuesta 3) Confirmar el envío de los datos. |
| Precondiciones | El link generado para la encuesta debe pertenecer a una replicación específica de un Usuario Investigador. |
| Post condiciones | La encuesta se registra en el área de trabajo del investigador específico de la replicación en curso. |
| Efectos Colaterales | Si el usuario no termina la encuesta, esta no se almacena en la base de datos. |
| Prioridad | Alta. |

Requisito funcional 5

El requisito funcional RR-RF005, denominado “Categorizar preguntas abiertas”, se encuentra detallado en la Tabla 21.

Tabla 21
Requisito funcional 5

| | |
|---------------------|--|
| Id. Requerimiento | RR-RF005 |
| Nombre | Categorizar preguntas abiertas. |
| Actor | Investigador Usuario, Investigador Administrador |
| Descripción | La aplicación web debe permitir categorizar las preguntas abiertas mediante la agrupación por categoría y subcategoría. |
| Entradas | Requisito: “Ejecutar Encuesta” con código: RR-RF004 |
| Salidas | - La categorización de cada pregunta es almacenada en la base de datos. 1) Ir al apartado de Replicaciones y Análisis de encuestas. 2) Seleccionar ver resultados. |
| Proceso | 3) Categorizar cada pregunta abierta del cuestionario. |
| Precondiciones | Cumplir con el requisito: “Ejecutar Encuesta” de código: RR-RF004. |
| Post condiciones | - La categorización de cada pregunta es almacenada en la base de datos. |
| Efectos Colaterales | En caso de surgir algún problema, se informa al usuario. |
| Prioridad | Alta. |

Requisito funcional 6

El requisito funcional RR-RF006, denominado “Analizar Encuestas”, se encuentra detallado en la Tabla 22.

Tabla 22
Requisito funcional 6

| | |
|-------------------|--|
| Id. Requerimiento | RR-RF006 |
| Nombre | Analizar Encuestas |
| Actor | Investigador Usuario, Investigador Administrador |
| Descripción | La aplicación web debe permitir seleccionar los parámetros deseados por el usuario para su posterior análisis. Los parámetros incluyen filtros de datos y selección de replicaciones a analizar. |
| Entradas | Requisito: “Ejecutar Encuesta” con código: RR-RF004 |
| Salidas | - Archivo LaTeX y PDF como informe de la replicación. 1) Ir al apartado de Replicaciones y Análisis de encuestas. 2) Configurar los parámetros deseados. 3) Enviar confirmación de análisis de encuestas. |
| Proceso | 4) Descargar los archivos PDF y LaTeX. |
| Precondiciones | Cumplir con el requisito: “Ejecutar Encuesta” de código: RR-RF004. |

CONTINÚA 

| | |
|---------------------|--|
| | -Los documentos generados se almacenan en la base de datos del sistema. |
| Post condiciones | - Los documentos generados se entregan al usuario mediante un archivo comprimido que contiene: documentos en LaTeX e imágenes del análisis de la encuesta. |
| Efectos Colaterales | En caso de surgir algún problema, se informa al usuario. |
| Prioridad | Alta. |

Requisito funcional 7

El requisito funcional RR-RF007, denominado “Respaldar documentos de investigación”, se encuentra detallado en la Tabla 23.

Tabla 23

Requisito funcional 7

| | |
|---------------------|--|
| Id. Requerimiento | RR-RF007 |
| Nombre | Respaldar documentos de investigación. |
| Actor | Investigador Usuario, Investigador Administrador |
| Descripción | La aplicación web debe permitir al usuario enviar un comprimido que contenga los documentos de LaTeX e imágenes utilizadas en el informe técnico. |
| Entradas | <ul style="list-style-type: none"> - Requisito: “Analizar Encuestas” con código: RR-RF006 - Comprimido con documentos de LaTeX e imágenes |
| Salidas | <ul style="list-style-type: none"> -Un mensaje de confirmación informa de operación exitosa. 1) Ejecutar el archivo .jar para iniciar el respaldo. 2) Ingresar a la aplicación web (include. RR-RF008). 3) La aplicación se encarga de subir automáticamente los archivos al servidor. |
| Proceso | |
| | Cumplir con el requisito: “Analizar Encuestas” de código: RR-RF006. |
| Precondiciones | Cumplir con el requisito: “Login de usuario” de código: RR-RF008. |
| Post condiciones | -El comprimido es almacenado en la base de datos. |
| Efectos Colaterales | La próxima vez que se vaya a descargar los archivos de la investigación se lo hará con la última versión que se encuentre alojada en el servidor. |
| Prioridad | Alta. |

Requisito funcional 8

El requisito funcional RR-RF008, denominado “Login de usuario”, se encuentra detallado en la Tabla 24.

Tabla 24
Requisito funcional 8

| | |
|---------------------|--|
| Id. Requerimiento | RR-RF008 |
| Nombre | Login de usuario |
| Actor | Investigador Usuario, Investigador Administrador |
| Descripción | La aplicación web debe permitir el ingreso al sistema de un usuario registrado. |
| Entradas | Nombre de usuario y contraseña. |
| Salidas | -Mensaje de ingreso exitoso, o de intento de nuevo, dependiendo del caso. |
| Proceso | <ol style="list-style-type: none"> 1) Ingreso a la pantalla de Login. 2) Llenado de usuario y contraseña. 3) Ingreso exitoso o usuario y/o contraseña incorrecta. |
| Precondiciones | - Cumplir con el requisito: “Registro de usuario investigador” de código: RR-RF009. |
| Post condiciones | - El usuario puede utilizar todas las funciones de la aplicación web según el rol concedido. |
| Efectos Colaterales | |
| Prioridad | Alta |

Requisito funcional 9

El requisito funcional RR-RF009, denominado “Solicitar creación de usuario investigador”, se encuentra detallado en la Tabla 25.

Tabla 25
Requisito funcional 9

| | |
|---------------------|---|
| Id. Requerimiento | RR-RF009 |
| Nombre | Solicitar creación de usuario investigador. |
| Actor | Usuario Externo. |
| Descripción | El aplicativo web debe permitir solicitar la creación de un usuario con el rol de “Investigador usuario”. |
| Entradas | Datos del usuario. |
| Salidas | -Mensaje de solicitud creada con éxito. |
| Proceso | <ol style="list-style-type: none"> 1) El usuario llena los datos necesarios para la creación de un nuevo usuario. 2) Se confirma el envío de la solicitud de creación de usuario. |
| Precondiciones | Ninguna. |
| Post condiciones | El usuario se almacena en la tabla <<USUARIOS>> de la base de datos listo para ser aprobado por el administrador de la aplicación web. |
| Efectos Colaterales | |
| Prioridad | Media |

Requisito funcional 10

El requisito funcional RR-RF010, denominado “Gestionar usuarios”, se encuentra detallado en la Tabla 26.

Tabla 26
Requisito funcional 10

| | |
|---------------------|---|
| Id. Requerimiento | RR-RF0010 |
| Nombre | Gestionar usuarios. |
| Actor | Investigador Administrador. |
| Descripción | El aplicativo web debe permitir al Investigador Administrador la gestión de usuarios. |
| Entradas | <ul style="list-style-type: none"> - Requisito: “Solicitar Creación de usuario investigador” con código: RR-RF009 |
| Salidas | <ul style="list-style-type: none"> - Mensaje de operación exitosa. - Mail de confirmación/notificación al usuario. |
| Proceso | <ol style="list-style-type: none"> 1) Ingresar a la aplicación web. 2) Acceder al apartado “Gestión de Usuarios”. 3) Realizar la actividad requerida: aprobación de usuario, modificación o eliminación. |
| Precondiciones | Ninguna. |
| Post condiciones | Se ha creado, modificado o eliminado un usuario. |
| Efectos Colaterales | |
| Prioridad | Media. |

Requisito funcional 11

El requisito funcional RR-RF011, denominado “Gestionar replications”, se encuentra detallado en la Tabla 27.

Tabla 27
Requisito funcional 11

| | |
|-------------------|---|
| Id. Requerimiento | RR-RF0011 |
| Nombre | Gestionar replications. |
| Actor | Investigador Administrador. |
| Descripción | El aplicativo web debe permitir al Investigador Administrador la gestión de replications. El actor puede aceptar o denegar permisos de replicación de los estudios que le pertenecen. |
| Entradas | <ul style="list-style-type: none"> - Requisito: “Solicitar Replicación de Estudio” con código: RR-RF003. |
| Salidas | <ul style="list-style-type: none"> - Mensaje de operación exitosa. - Mail de confirmación/notificación al usuario. |

| | |
|---------------------|--|
| Proceso | <ol style="list-style-type: none"> 1) Ingresar a la aplicación web. 2) Acceder al apartado “Gestión de replicaciones”. 3) Realizar la actividad requerida: aprobación de replicación, modificación o eliminación. |
| Precondiciones | Ninguna. |
| Post condiciones | Se ha creado, modificado o eliminado una replicación. |
| Efectos Colaterales | Permitir o denegar el permiso de replicación a un investigador usuario. |
| Prioridad | Alta. |

4.3.1.2.Requisitos no funcionales

Requisito no funcional 1

El requisito no funcional RR-RNF001, descrito en la Tabla 28, define restricciones acerca de la interfaz de usuario.

Tabla 28

Requisito no funcional 1

| | |
|-------------------|--|
| Id. Requerimiento | RR-RNF001 |
| Nombre | Acerca de la interfaz |
| Descripción | <ul style="list-style-type: none"> • La plataforma web tendrá una interfaz accesible e intuitiva, de manera que los procesos que debe llevar el usuario serán muy claros. • La plataforma web tendrá su interfaz de usuario en idioma español. |

Requisito no funcional 2

El requisito no funcional RR-RNF002, descrito en la Tabla 29, describe qué características restrictivas va a tener la navegación de la plataforma web.

Tabla 29

Requisito no funcional 2

| | |
|-------------------|--|
| Id. Requerimiento | RR-RNF002 |
| Nombre | Acerca de la navegación |
| Descripción | <ul style="list-style-type: none"> • La plataforma web tendrá consistencia de visualización entre los diferentes navegadores. • La plataforma web tendrá medios de navegación estándar (inicio, regresar, salir, atrás, guardar, enviar, etc.). • La plataforma web tendrá diseño responsivo. |

Requisito no funcional 3

El requisito no funcional RR-RNF003, descrito en la Tabla 30, describe las características de la usabilidad dentro de la plataforma web.

Tabla 30

Requisito no funcional 3

| | |
|-------------------|---|
| Id. Requerimiento | RR-RNF003 |
| Nombre | Acerca de la usabilidad <ul style="list-style-type: none"> • La plataforma web contará con una interfaz atractiva, simple y de fácil uso. |
| Descripción | <ul style="list-style-type: none"> • La plataforma web mostrará mensajes informativos constantemente a fin de tener buena comunicación con el usuario. |

Requisito no funcional 4

El requisito no funcional RR-RNF004, descrito en la Tabla 31, especifica las condiciones que debe cumplir la plataforma web con respecto a la seguridad del sistema.

Tabla 31

Requisito no funcional 4

| | |
|-------------------|--|
| Id. Requerimiento | RR-RNF004 |
| Nombre | Acerca de la seguridad <ul style="list-style-type: none"> • La plataforma web almacenará de manera íntegra la información almacenada. |
| Descripción | <ul style="list-style-type: none"> • La plataforma web protegerá el acceso no autorizado a los datos almacenados dentro de ella. |

4.3.2 Casos de uso

Con la especificación de requisitos se realizó el diagrama de casos de uso, donde se observan las interacciones entre actores y casos de uso

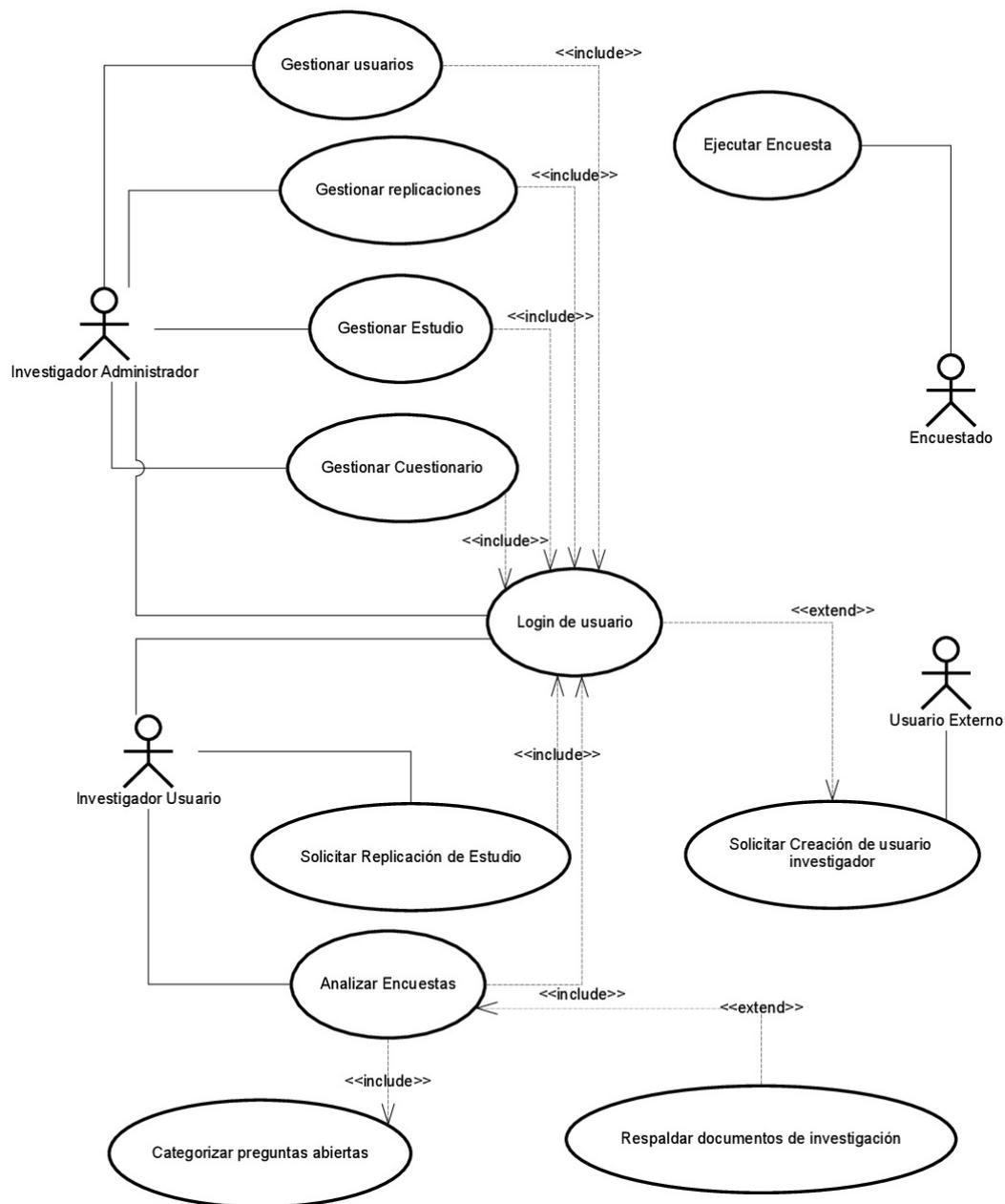


Figura 17. Diagrama de casos de uso

4.4 ARQUITECTURA DE SOFTWARE

Con el fin de ilustrar la arquitectura de software que se utilizó en el funcionamiento de la plataforma web se tiene la Figura 18.

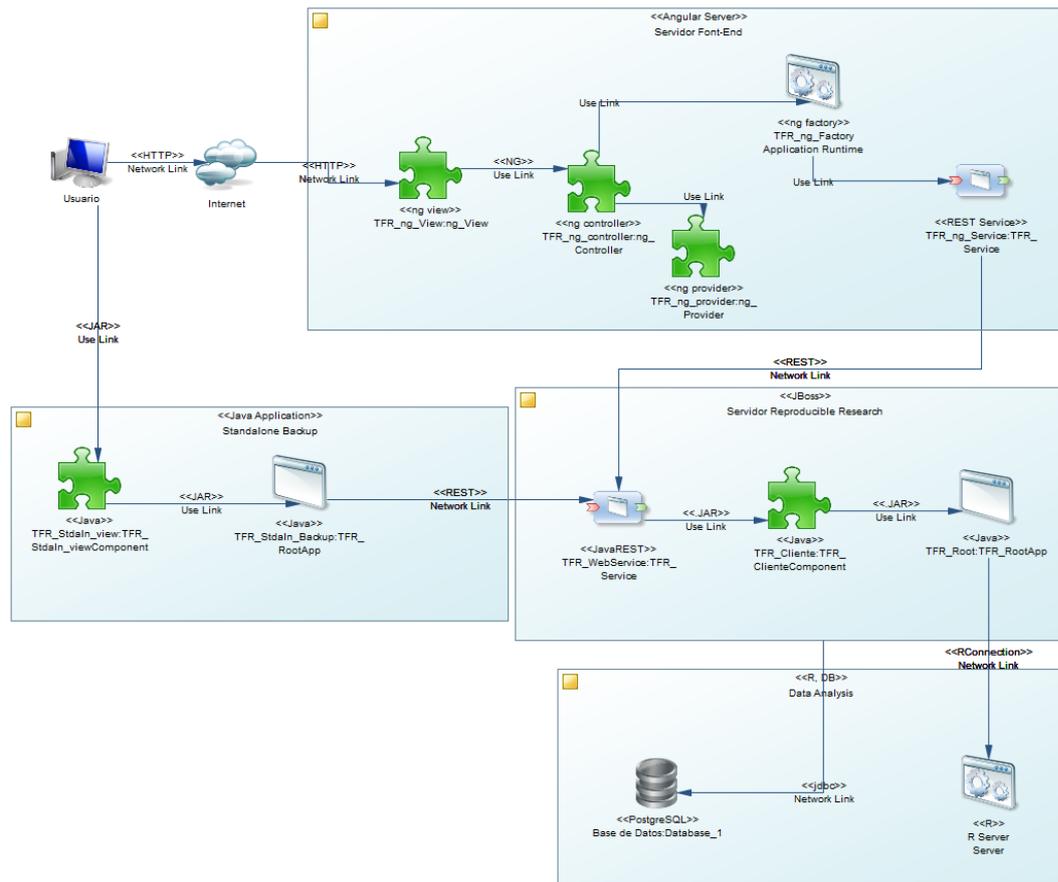


Figura 18. Arquitectura de software

Toda la plataforma web está diseñada en cuatro zonas muy importantes. Estas cuatro zonas se agrupan en dos programas, la plataforma web en sí y un aplicativo standalone desarrollado para Java SE Runtime Environment. A continuación, se detalla el funcionamiento de cada programa.

4.4.1. Arquitectura plataforma web

La arquitectura de la plataforma web se encuentra aislada en la Figura 19. Consta de

tres módulos: Frontend con Angular 7, Backend con Java 8 y análisis de datos. Como ente consumidor de este programa está un usuario, quien a través del internet se conecta al frontend de la plataforma web.

Angular 7

El primer componente se encarga totalmente del despliegue de información para establecer la interacción entre el usuario y el sistema. Este proyecto desarrollado en Angular 7 consta de cinco componentes. Todo inicia con el componente “TFR_ng_View” el cual recibe las peticiones del usuario y muestra las pantallas en formato HTML5, este a su vez comunica las peticiones al componente “TFR_ng_controller” el cual las gestiona y las conduce a los dos componentes aledaños. “TFR_ng_provider” se encarga de codificar la parte visual utilizando una combinación de código HTML5 y CCS, es así que dibuja botones, ventanas, cuadros de texto, mensajes emergentes y otros componentes visuales. “TFR_ng_Factory”, por otro lado, maneja todos los objetos, entidades, clases y datos que van a ser representados en pantalla. Finalmente, las peticiones terminan en el último componente denominado “TFR_ng_Service”, el cual es la puerta de comunicación con el módulo backend, y se encarga de gestionar y consumir servicios mediante el protocolo REST.

Java 8

El módulo backend desarrollado en Java está programado en tres capas. Empezando por la capa más externa se tiene a “TFR_WebService”, un componente de tipo JavaREST encargado de poner a disposición a todos los servicios REST, los servicios ofrecidos por este módulo satisfacen todos los requisitos e historias de usuario.

Luego se tiene la capa “TFR_Cliente” netamente encargada de ejecutar la lógica del negocio, es el componente donde se procesan las peticiones del usuario para la gestión de la información dentro de la plataforma web. Finalmente, en la capa más interna está “TFR_Root”,

una colección de interfaces y donde se establece el vínculo con el último módulo del sistema, también está encargada de comunicarse con la base de datos y de enviar las peticiones de análisis de datos al servidor R.

Base de datos y R

El último módulo en esta arquitectura está destinado al procesamiento y análisis de datos. Dentro de este se tiene un servidor de base de datos PostgreSQL y un servidor de R. Las peticiones de gestión de información son atendidas por la base de datos con ayuda del gestor de base de datos, mientras que las peticiones para analizar estadísticamente la información y generar gráficos de resultados las atiende el servidor de R.

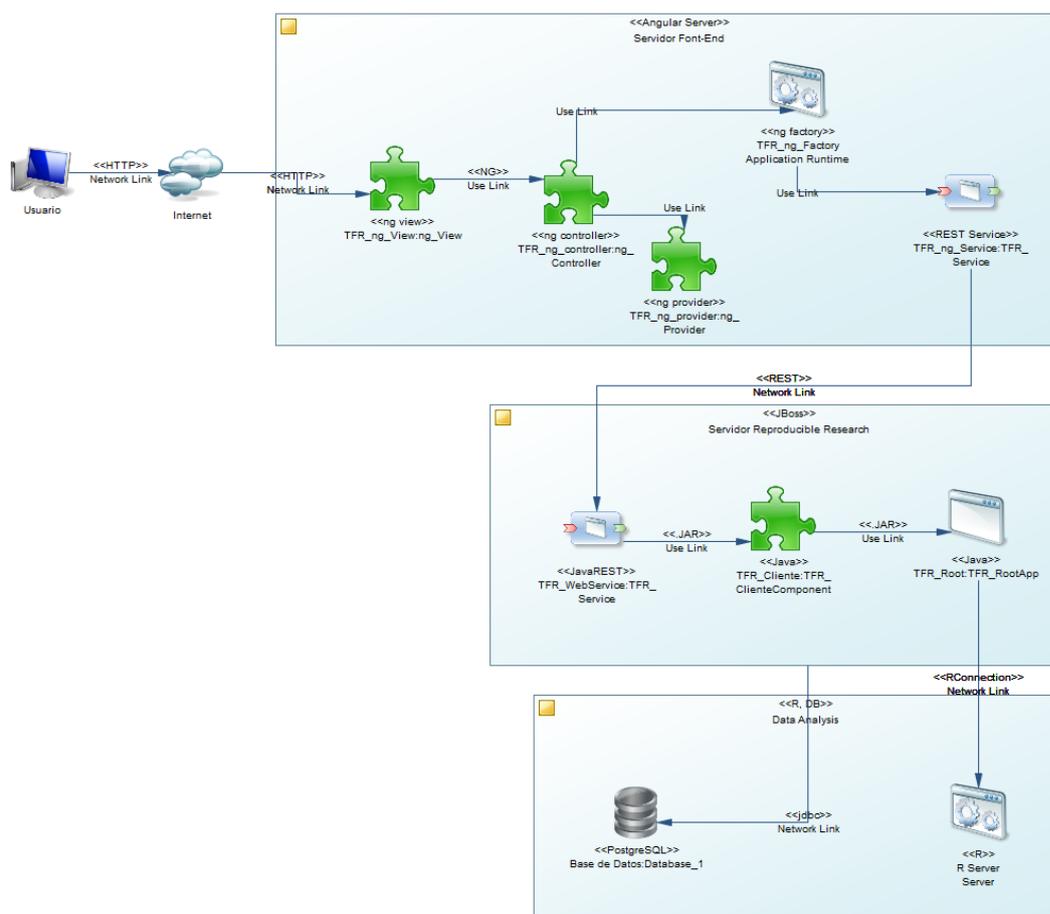


Figura 19. Arquitectura plataforma web

4.4.2. Arquitectura programa standalone

El programa standalone fue diseñado para cumplir un requisito funcional en específico, el RR-RF007 (ver

Tabla 7). Esta funcionalidad se la puede ejecutar desde la plataforma web también, pero al utilizar el programa standalone se tiene mayor facilidad para respaldar los archivos de resultados en la nube. Funciona como un programa ejecutable desarrollado para la máquina virtual de Java.

Dentro de la arquitectura, descrita en la Figura 20, se observa que tiene dos módulos. “TFR_Stdaln_view” corresponde en su totalidad al aspecto visual de la aplicación, permitiendo la interacción del usuario con el programa. Mientras que “TFR_Stdaln_Backup” lleva a cabo todos algoritmos para comunicarse con el servidor y consumir los servicios REST que sean pertinentes. Los dos módulos están programados netamente en Java 8.

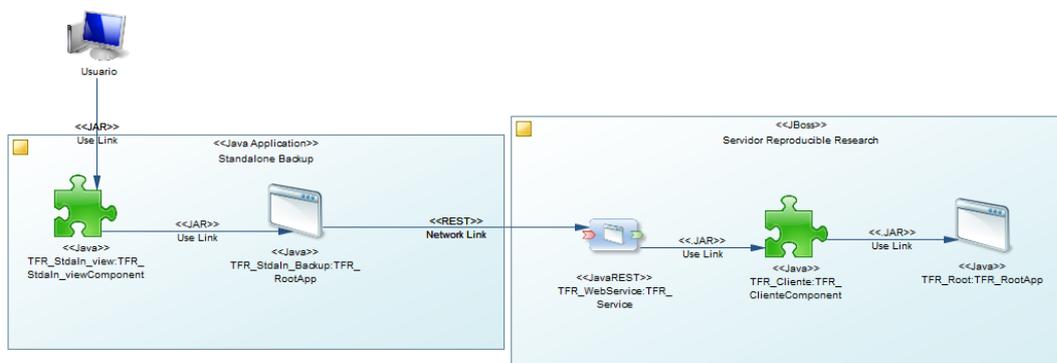


Figura 20. Arquitectura programa standalone

4.5 INTERFAZ DE USUARIO

La interfaz de usuario de la herramienta se la presenta en ocho pantallas, las cuales se describen a continuación. Algunas funcionalidades de la plataforma se han organizado visualmente mediante pestañas de navegación.

Login

La interfaz de usuario destinada a la funcionalidad de Login se muestra en la Figura 21. Consta de un menú principal donde se despliegan todos los estudios publicados en la plataforma web, es decir, todos los estudios que están disponibles para ser replicados, además se puede conocer un poco más sobre cada estudio para aumentar el interés de replicación. Por la barra superior se encuentra el login donde se puede ingresar al sistema o solicitar la creación de un usuario.

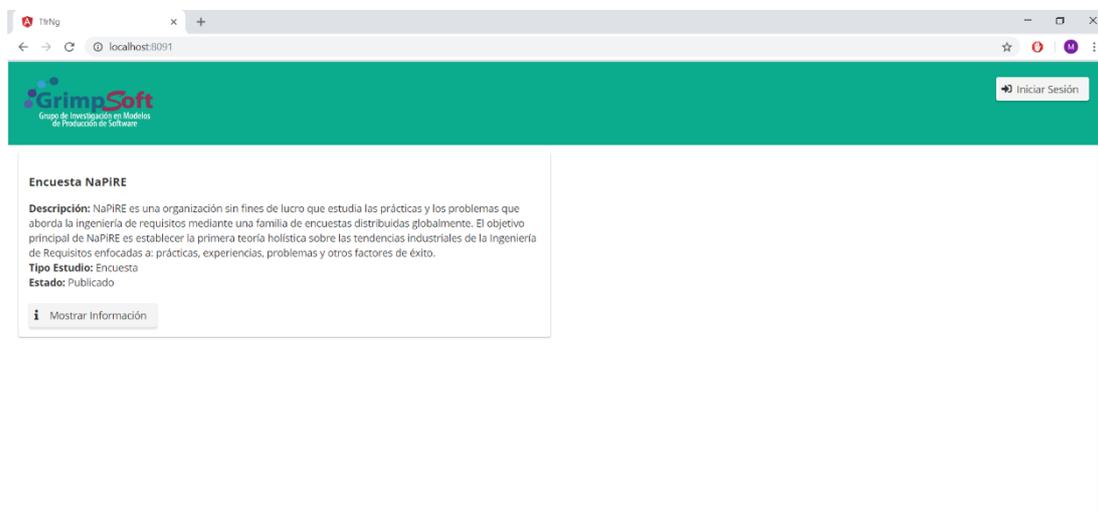


Figura 21. Interfaz de usuario Login

Gestión de solicitudes

La interfaz de usuario de la Figura 22 está destinada a la aprobación o denegación de las solicitudes de creación de usuario y de los permisos para replicar estudios. En este apartado también se puede solicitar más información al usuario en caso de que no cumpla los requisitos mínimos para aceptar cualquiera de los dos tipos de solicitud.

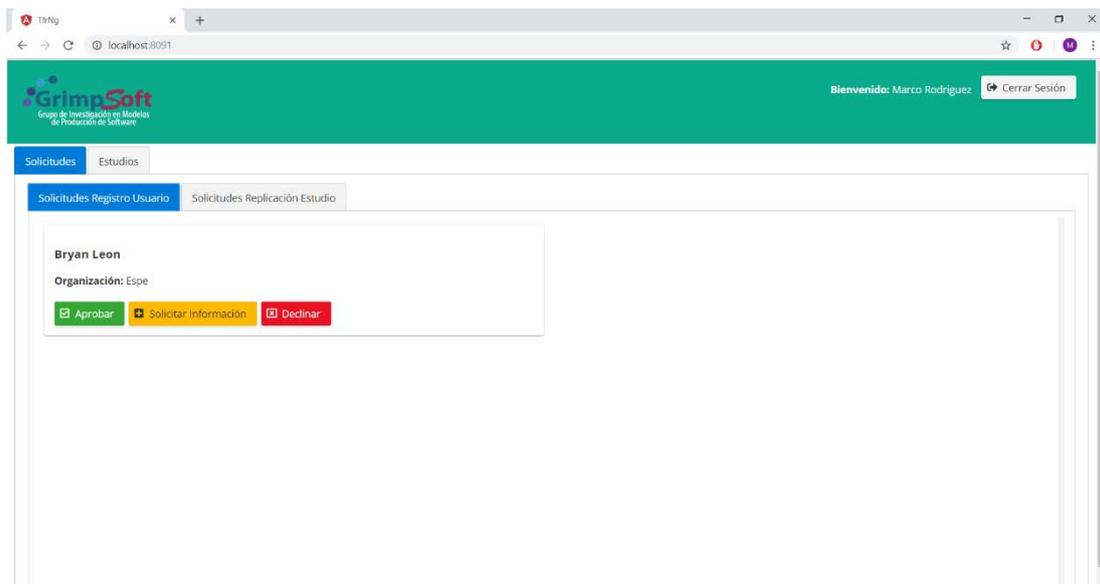


Figura 22. Interfaz de usuario gestión de solicitudes

Gestión de estudios

La interfaz de la Figura 23 muestra la pantalla donde se exponen los estudios para ser replicados, visto desde la perspectiva de un usuario que ha ingresado a la plataforma. Aquí el usuario puede seleccionar el estudio que desea replicar, de esta forma se envía la solicitud de replicación al administrador del estudio. Otra opción permite ver las replicaciones que tiene el estudio seleccionado, mientras que el botón “administrar archivos” sirve para respaldar en la nube el reporte de resultados.

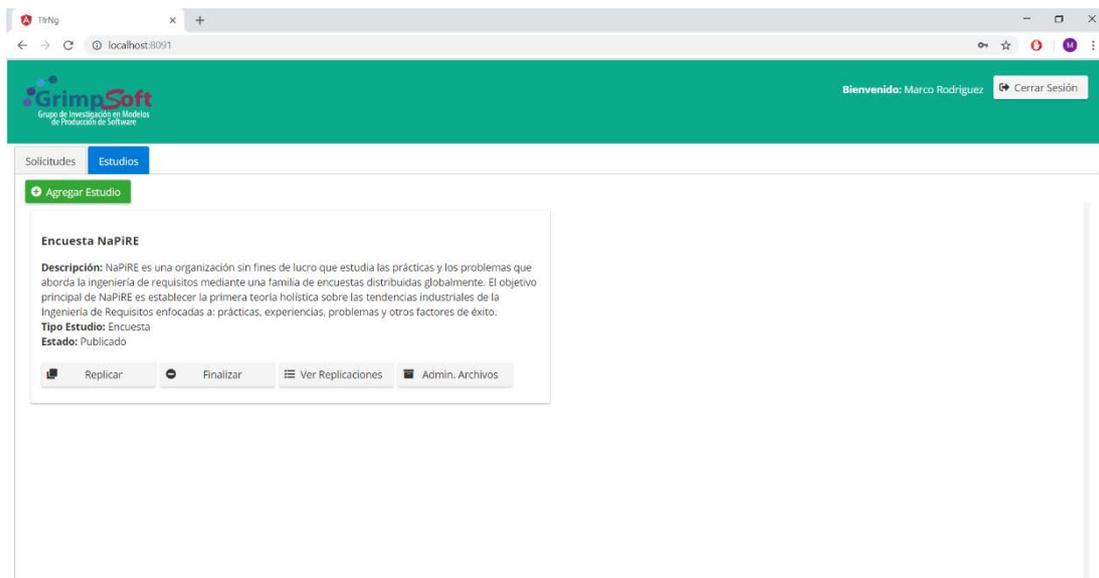


Figura 23. Interfaz de usuario gestión de estudios

Como facultades exclusivas del administrador, se tiene la pantalla de creación de estudios, misma que se muestra en la Figura 24, además de poder finalizar un estudio, es decir, que ya no se encuentre disponible para ser replicado.

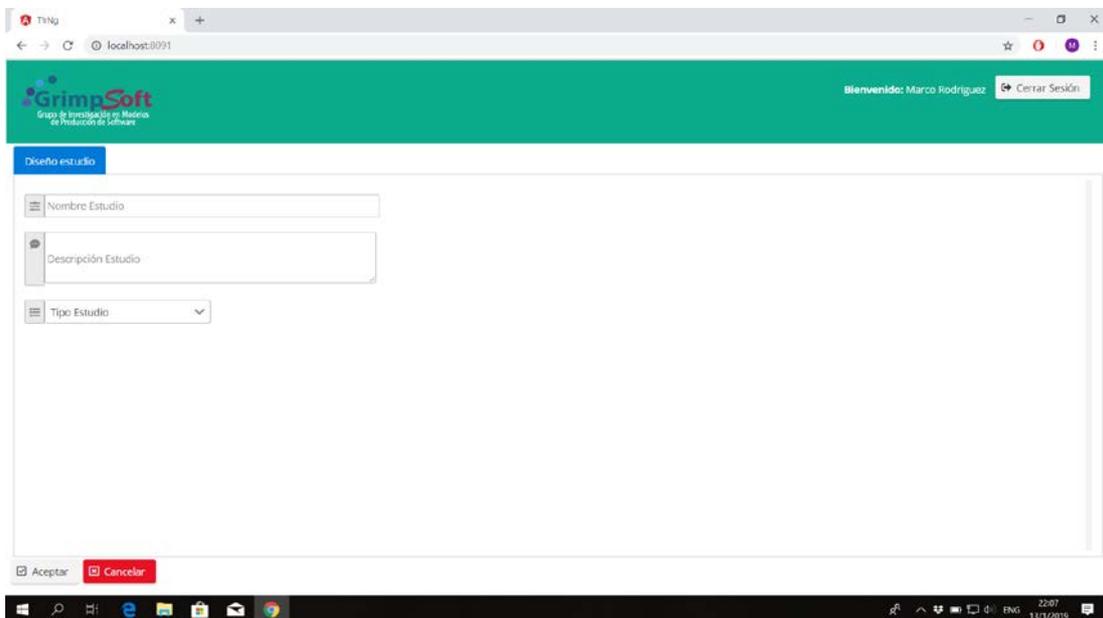


Figura 24. Interfaz de usuario diseño de estudio

Replicaciones

En la pantalla representada por la Figura 25 se observa la interfaz que permite interactuar con las replicaciones de los estudios. Aquí se muestran todas las replicaciones de un estudio en particular. Este es el punto de partida para el análisis de los resultados de las encuestas, también permite difundir las encuestas de la replicación seleccionada o darla por finalizada.

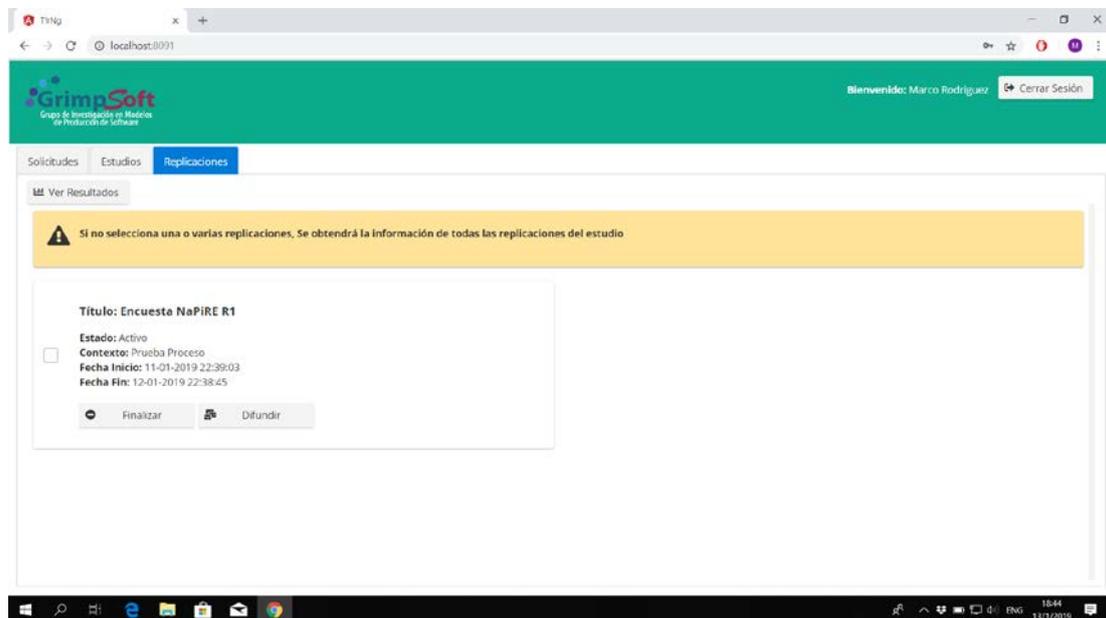


Figura 25. Interfaz de usuario replicaciones

Antes de comenzar con el análisis de datos es necesario categorizar las respuestas a las preguntas abiertas del cuestionario, para ello se utiliza la interfaz de la Figura 26. Aquí se muestran todas las respuestas agrupadas por preguntas, de modo que se selecciona categoría y subcategoría para cada respuesta.

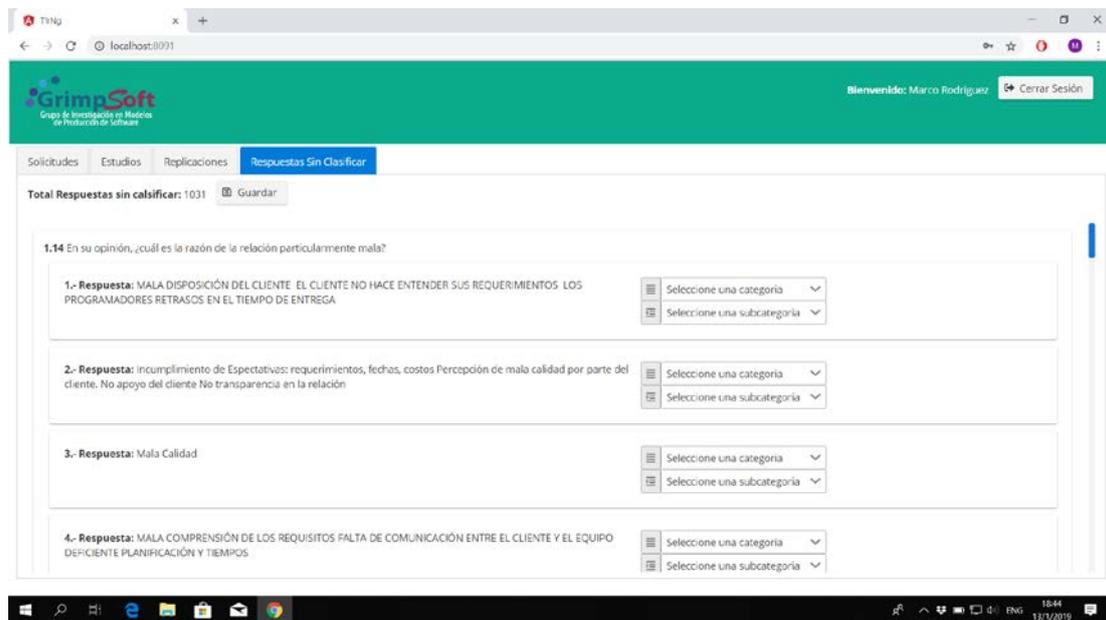


Figura 26. Interfaz de usuario categorizar respuestas

Resultados

La interfaz de la Figura 27 ilustra la pantalla de resultados. Aquí se visualizan mediante gráficos de barra o pastel los resultados a cada pregunta de la encuesta, además permite descargar un archivo comprimido que contiene las gráficas presentadas y la plantilla del reporte de resultados.

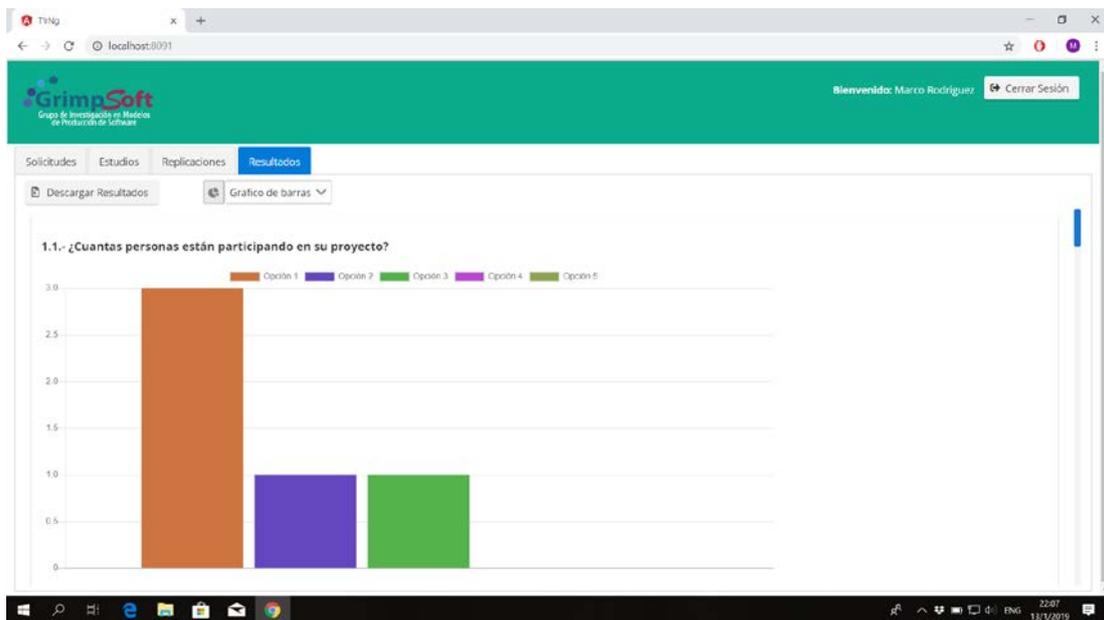


Figura 27. Interfaz de usuario resultados

Encuesta

La interfaz presentada en la Figura 28 muestra la pantalla donde el encuestado va a ingresar las respuestas a las preguntas del cuestionario. Está organizado por pestañas para una navegación más cómoda.

Encuesta

Datos encuestado | Datos del proyecto | Elicción de requisitos | Documentación | Validación de requisitos | Gestión de requisitos | Problemas, Causas y Efectos | Información Adicional

Información de la organización

Género: Seleccione un género

País: Seleccione un país

Provincia: Seleccione una provincia

Ciudad: Seleccione una ciudad

Finalizar

Figura 28. Interfaz de usuario encuesta

CAPÍTULO V

APLICACIÓN DEL PROTOTIPO AL CONTEXTO/PROBLEMA

5.1 CASO DE ESTUDIO

El prototipo propuesto utiliza al proyecto NaPiRE como caso de estudio con el propósito de validar la herramienta y evaluar la eficacia que ofrece al aplicar el paradigma de investigación reproducible dentro de la ingeniería de software.

5.1.1. Proyecto NaPiRE

Naming the Pain in Requirements Engineering (NaPiRE) es una organización sin fines de lucro que estudia las prácticas y los problemas que aborda la ingeniería de requisitos mediante una familia de encuestas distribuidas globalmente, estas encuestas se ejecutan dos veces al año. El objetivo principal de NaPiRE es establecer la primera teoría holística sobre las tendencias industriales de la IR enfocadas a dar respuesta a las siguientes interrogantes:

- ¿Cuáles son las prácticas y experiencias de la IR en la industria?
- ¿Cómo se manifiestan los problemas en el proceso de IR?
- ¿Cuáles son los posibles factores de éxito en la IR?

Esta organización está conformada por un gran número de países miembros, los cuales contribuyen al logro de los objetivos planteados.

5.1.2. Descripción de la encuesta NaPiRE

La formulación de la encuesta NaPiRE se llevó a cabo para que participen todos los miembros involucrados en el proyecto, con preferencia a los profesionales que se dedican al proceso de ingeniería de requisitos. La encuesta está compuesta por 40 preguntas, las cuales están dividida en 6 diferentes categorías.

La encuesta está formulada con preguntas tanto abiertas como cerradas, las opciones de respuesta de las preguntas cerradas pueden ser de dos tipos: opción única o múltiple;

adicionalmente se incluye la opción “otro” con el fin de obtener desviaciones específicas de la encuesta.

Algunas preguntas necesitan medir el grado de satisfacción para lo cual la encuesta utiliza la escala Likert que mide las actitudes o grado de conformidad. Esta escala tiene cinco niveles de respuesta, indicando el grado de acuerdo o desacuerdo de la persona encuestada.

5.2 VALIDACIÓN DE LA PROPUESTA

La Ingeniería de Software ha tenido un aumento significativo en el número de estudios empíricos en las últimas décadas (Larsson et al., 2014). Mientras los estudios empíricos se hacen más comunes en la Ingeniería de Software, la importancia de que estos sean reproducibles aumenta (Larsson et al., 2014). El problema surge al intentar sintetizar esta gran cantidad de trabajos ya que puede ser una tarea ardua, principalmente por la dificultad que implica la verificación de los resultados de las investigaciones existentes (Fernández, 2011; Sjöberg et al., 2005), por otra parte, replicar estudios puede llegar a ser muy agotador si no se los ha diseñado con este propósito debido a que no se publican detalles profundos utilizados en la investigación (Fomel, 2015). Los beneficios de utilizar el paradigma de investigación reproducible repercuten en que las repeticiones ayudan a aumentar el cuerpo de conocimiento alrededor de Ingeniería de Software, lo que a su vez conduce a una mayor madurez del campo. Una replicación de un estudio también viene con beneficios para el estudio original en términos de aumento de la confianza para el estudio conducido y los resultados informados. Es por esto que con el motivo de integrar el paradigma de investigación reproducible a la ingeniería de software se desarrolló el prototipo de investigación reproducible aplicado al caso de estudio NaPiRE, el cual busca facilitar la replicación de investigaciones de tipo encuesta. Para la evaluación de prototipo se aplicó el siguiente diseño de encuesta de opinión.

5.2.1. Diseño de la encuesta

El diseño de la encuesta se basó en las fases de diseño de encuestas propuestas por J. Linares (Linares Fontela, 2015). Dentro de las fases propuestas se encuentra la fase “Diseñar cuestionario”, la cual consta de las siguientes actividades: definir el alcance de la encuesta, determinar las preguntas y definir el tipo de formato de las preguntas. A estas actividades se les añadió ciertos elementos para mejorar su descripción. Las actividades del diseño se desglosan a continuación a manera de subtítulos.

Objetivo de la encuesta

El objetivo principal es evaluar la utilidad del prototipo de investigación reproducible aplicado a los estudios tipo survey en la ingeniería de software mediante la aplicación de la encuesta de satisfacción de usuario.

Preguntas de investigación

Con el motivo de definir el alcance de la encuesta se plantearon las siguientes preguntas de investigación.

RQ1. ¿Es posible desarrollar una herramienta que ayude a aplicar investigación reproducible en estudios de ingeniería de software?

RQ2. ¿Qué fases de la investigación tradicional son más fácilmente adaptables al paradigma de investigación reproducible?

RQ3. ¿Qué beneficios aporta la investigación reproducible a la ingeniería de software?

Población objetivo

De acuerdo con (Hernández Sampieri, Fernández Collado, & Baptista Lucio, 2010) se deben definir las características de las personas a quienes va dirigida la encuesta, es así que se las detalla a continuación.

Características

- **Género:** Sin distinción.
- **Instrucción académica:** pregrado o postgrado.
- **Edad:** mayor a 18 años.

Requisitos

- Tener conocimientos en la aplicación del método científico en ingeniería de software.
- Investigador con experiencia básica en la ejecución de estudios empíricos.

Correspondencia de preguntas

En la Tabla 32 se muestra la correlación que tienen las preguntas de investigación con las preguntas del cuestionario.

Tabla 32

Correspondencia de preguntas

| PREGUNTAS DE INVESTIGACIÓN | CUESTIONARIO |
|--|--|
| <p>RQ1. ¿Es posible desarrollar una herramienta que ayude a aplicar investigación reproducible en estudios de ingeniería de software?</p> | <p>1. ¿Considera usted que la herramienta propuesta facilita la aplicación del paradigma de investigación reproducible al caso de estudio “NaPiRE”?</p> |
| <p>RQ2. ¿Qué fases de la investigación tradicional son más fácilmente adaptables al paradigma de investigación reproducible?</p> | <p>2. De las siguientes fases dentro de la investigación tradicional, seleccione los que considera que han sido agilizados gracias al prototipo de investigación reproducible. Considere como referencia la replicación de un estudio existente, mas no la creación de un nuevo estudio.</p> <ol style="list-style-type: none"> Definición del problema Hipótesis Razonamiento y deducción Observación, prueba y experimento Resultados |

CONTINÚA 

RQ3. ¿Qué beneficios aporta la investigación reproducible a la ingeniería de software?

3. ¿Cómo le ayudó la herramienta propuesta a replicar el caso de estudio “NaPiRE”?
 4. Mediante la escala presentada, evalúe los beneficios de la investigación reproducible que encontró en la herramienta propuesta.
 - a. La herramienta le ayuda a entender el proceso completo seguido durante la investigación.
 - b. La herramienta proporciona los archivos utilizados en la investigación.
 - c. La herramienta muestra el diseño de la encuesta utilizada en el estudio.
 - d. La herramienta aumenta la confianza en los resultados del caso de estudio.
 - e. La herramienta ayuda a contribuir con nuevos resultados apegados al marco común del caso de estudio principal.
-

5.2.1.1. Encuesta de satisfacción de usuario

Introducción

La presente encuesta tiene como propósito evaluar la utilidad del prototipo de investigación reproducible aplicado al caso de estudio NaPiRE en la ingeniería de software.

El paradigma de investigación reproducible es una alternativa para realizar investigación científica. Su principal cualidad es que muestra los métodos, protocolos, técnicas, herramientas, procesos y recursos que fueron utilizados durante la investigación, de manera que otro investigador pueda replicar toda la investigación en una nueva instancia propia; de esta forma, los resultados pueden validarse y aumentar su nivel de veracidad (Larsson et al., 2014).

Una replicación de un estudio también viene con beneficios para el estudio original en términos de aumento de la confianza para el experimento conducido y los resultados informados (por ejemplo, los intervalos de confianza más estrictos) (Larsson et al., 2014).

NaPiRE es una organización sin fines de lucro que estudia las prácticas y los problemas que aborda la ingeniería de requisitos mediante una familia de encuestas distribuidas globalmente (Méndez et al., 2012). El objetivo principal de NaPiRE es establecer la primera teoría holística sobre las tendencias industriales de la Ingeniería de Requisitos enfocadas en: prácticas, experiencias, problemas y otros factores de éxito (Méndez et al., 2012).

Se eligió al proyecto NaPiRE como caso de estudio con el fin de adaptar el paradigma de investigación reproducible a la ingeniería de software. Todas las preguntas del cuestionario NaPiRE se introdujeron a la herramienta propuesta para ser parte del proceso de replicación. El prototipo propuesto guía al investigador interesado en replicar este caso de estudio desde la aplicación de la encuesta, hasta el análisis de resultados.

Para la evaluación del prototipo es necesario que usted imagine que desea utilizar la encuesta propuesta por NaPiRE en una determinada población (por ejemplo, solo la ciudad de Guayaquil), con el fin de conocer los problemas y prácticas que tienen las empresas desarrolladoras de software en dicho lugar. Para ello, usted va a planificar y empezar su propia investigación, por lo que tiene dos opciones, empezar la investigación con sus propios medios de manera tradicional o utilizar el prototipo propuesto basado en el paradigma de investigación reproducible. Con ayuda de la herramienta propuesta, va a realizar una replicación de la investigación NaPiRE, posterior a esto le pedimos de favor que conteste las preguntas presentadas a continuación.

Datos del investigador encuestado

Universidad a la que pertenece: _____

Cargo: _____

Cuestionario

1. ¿Considera usted que la herramienta propuesta facilita la aplicación del paradigma de investigación reproducible al caso de estudio “NaPiRE”? Justifique su respuesta.

2. De las siguientes fases dentro de la investigación tradicional, encierre en un círculo el literal que considera que han sido agilizados gracias al prototipo de investigación reproducible. Considere como referencia la replicación de un estudio existente, mas no la creación de un nuevo estudio.

- a. Definición del problema
- b. Hipótesis
- c. Razonamiento y deducción
- d. Observación, prueba y experimento
- e. Resultados

3. ¿Cómo le ayudó la herramienta propuesta a replicar el caso de estudio “NaPiRE”?

4. Mediante la escala presentada, evalúe los beneficios de la investigación reproducible que encontró en la herramienta propuesta.

- a. La herramienta le ayuda a entender el proceso completo seguido durante la investigación.
- Muy de acuerdo
 - De acuerdo
 - Indiferente
 - Desacuerdo
 - Muy en desacuerdo
- b. La herramienta proporciona los archivos utilizados en la investigación.
- Muy de acuerdo
 - De acuerdo
 - Indiferente
 - Desacuerdo
 - Muy en desacuerdo
- c. La herramienta muestra el diseño de la encuesta utilizada en el estudio.
- Muy de acuerdo
 - De acuerdo
 - Indiferente
 - Desacuerdo
 - Muy en desacuerdo
- d. La herramienta aumenta la confianza en los resultados del caso de estudio
- Muy de acuerdo
 - De acuerdo

- Indiferente
 - Desacuerdo
 - Muy en desacuerdo
- e. La herramienta ayuda a contribuir con nuevos resultados apegados a al marco común del caso de estudio principal
- Muy de acuerdo
 - De acuerdo
 - Indiferente
 - Desacuerdo
 - Muy en desacuerdo

5.2.2. Proceso de ejecución de encuestas

El proceso de ejecución de las encuestas constó de dos fases. En una primera fase se instruyó a los investigadores la forma en que debían utilizar la herramienta y en qué consistía la demostración de la misma, las instrucciones de uso de la herramienta se encuentran en la sección 5.2.2.1. La segunda fase consistió en que los investigadores contesten la encuesta en cuestión, la misma que tiene el análisis de resultados en la sección 5.2.3.

5.2.2.1. Instrucciones de uso de la herramienta

El presente documento tiene la finalidad de guiarle en el proceso de replicación del caso de estudio NaPiRE a través de la herramienta de investigación reproducible que le ha sido entregada. Tomando como punto de partida el hecho que se ha entendido el contexto en el cual usted va a utilizar la herramienta se procede a explicar paso a paso las actividades que se le solicita que realice a fin de evaluar el prototipo propuesto, además se aclara que algunas actividades fueron obviadas, mientras que otras van a tener asistencia para agilizarlas.

Actividad de login

Como primer paso usted va a ingresar a la URL: grimpsoft.espe.edu.ec/TFR/. Aquí encontrará la pantalla de login (ver Figura 29) donde va a ingresar al sistema con las credenciales que le han sido entregadas con el fin de ingresar a las funcionalidades del sistema.

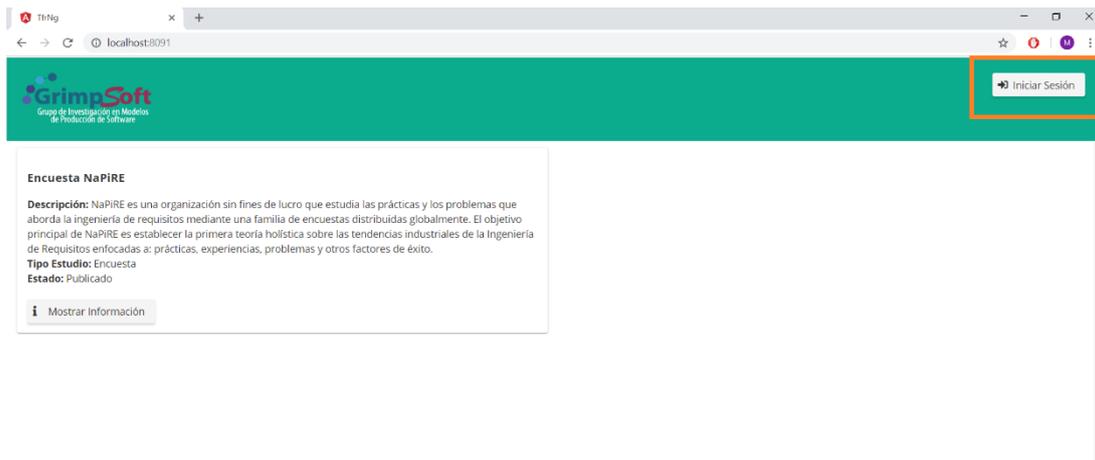


Figura 29. Actividad de login

Actividad elegir estudio para replicar

En el listado de estudios disponibles para replicar está el caso de estudio NaPiRE. Como muestra la Figura 30 se debe dar clic en el botón “Replicar” con el fin de iniciar una replicación del estudio, luego debe esperar unos minutos hasta que se le otorgue permiso de replicación y usted deberá dar clic en el botón “Ver Replicaciones” para empezar con las actividades de replicación.

The screenshot shows the GrimpSoft interface. At the top, there is a green header with the logo and text: "GrimpSoft Grupo de Investigación en Modelos de Producción de Software". Below the header, there are two tabs: "Solicitudes" and "Estudios", with "Estudios" being the active tab. A green button labeled "+ Agregar Estudio" is visible. The main content area displays a study titled "Encuesta NaPiRE". The study description states: "Descripción: NaPiRE es una organización sin fines de lucro que estudia las prácticas y los problemas que aborda la ingeniería de requisitos mediante una familia de encuestas distribuidas globalmente. El objetivo principal de NaPiRE es establecer la primera teoría holística sobre las tendencias industriales de la Ingeniería de Requisitos enfocadas a: prácticas, experiencias, problemas y otros factores de éxito." The study type is "Tipo Estudio: Encuesta" and the status is "Estado: Publicado". At the bottom of the study card, there are four buttons: "Replicar" (highlighted with an orange border), "Finalizar", "Ver Replicaciones", and "Admin. Archivos".

Figura 30. Actividad elegir estudio para replicar

Actividad de difusión

Como muestra la Figura 31, debe dar clic en el botón “Difundir”, el mismo que le solicitará las direcciones de correo electrónico a los cuales se les enviará el mail con la invitación a participar de la encuesta NaPiRE, después le pedirá el diseño de la imagen de invitación (flyer), en donde, a partir de una plantilla usted podrá aumentar o quitar texto a la invitación.

GrimpSoft
Grupo de Investigación en Modelos de Producción de Software

Solicitudes Estudios **Replicaciones**

Ver Resultados

! Si no selecciona una o varias replicaciones, Se obtendrá la información de todas las replicaciones

Título: Encuesta NaPIRE R1

Estado: Activo

Contexto: Prueba Proceso

Fecha Inicio: 11-01-2019 22:39:03

Fecha Fin: 12-01-2019 22:38:45

Finalizar **Difundir**

Figura 31. Actividad de difusión

Actividad de análisis de resultados

A fin de agilizar el proceso se han completado algunas encuestas, de manera que usted ya las tiene categorizadas y listas para su análisis. La pantalla presentada en Figura 32 muestra los resultados de la encuesta. En este punto usted va a dar clic al botón “Descargar Resultados” y la plataforma web va a empezar a empaquetar todos los resultados y entregarle en una plantilla de reporte de resultados. Después usted debería desarrollar todo el reporte de resultados siguiendo la estructura entregada, pero por fines didácticos no tiene que hacerlo.

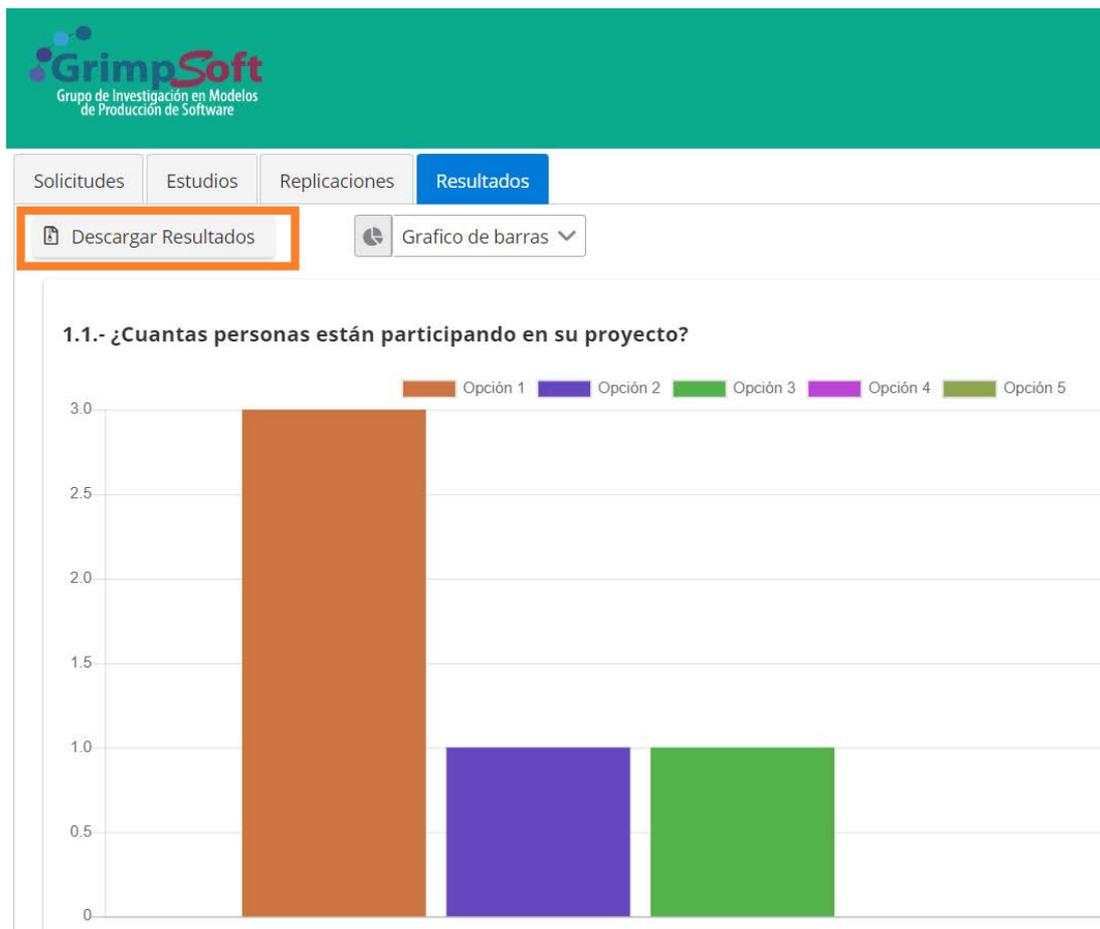


Figura 32. Actividad de análisis de resultados

Actividad respaldar resultados

Una vez que ha concluido el reporte de resultados se procede a respaldar los archivos finales. Para ello tiene dos opciones: en el menú de estudios (ver Figura 33) debe dar clic en el botón “Admin. Archivos” y subir el comprimido zip a la plataforma web o utilizar el programa standalone incluido en el comprimido de resultados que le fue entregado en la actividad anterior, simplemente da doble clic al ejecutable “backup.jar” (ver Figura 34), ingresa sus credenciales y espera a que automáticamente se envíen los archivos al servidor.

GrimpSoft
Grupo de Investigación en Modelos de Producción de Software

Solicitudes | Estudios

+ Agregar Estudio

Encuesta NaPiRE

Descripción: NaPiRE es una organización sin fines de lucro que estudia las prácticas y los problemas que aborda la ingeniería de requisitos mediante una familia de encuestas distribuidas globalmente. El objetivo principal de NaPiRE es establecer la primera teoría holística sobre las tendencias industriales de la Ingeniería de Requisitos enfocadas a: prácticas, experiencias, problemas y otros factores de éxito.

Tipo Estudio: Encuesta

Estado: Publicado

Replicar Finalizar Ver Replicaciones **Admin. Archivos**

Figura 33. Actividad respaldar resultados desde plataforma

resultados.zip

File Commands Tools Favorites Options Help

Add Extract To Test View Delete Find Wizard Info VirusScan Comment SFX

↑ resultados.zip - ZIP archive, unpacked size 11.160.807 bytes

| Name | Size | Packed | Type | Modified | CRC32 |
|------------------|-----------|-----------|---------------------|-----------------|----------|
| .. | | | Carpeta de archivos | | |
| images | | | Carpeta de archivos | | |
| backup.jar | 6.159.449 | 5.547.376 | Executable Jar File | 17/01/2019 5:08 | 335381AC |
| paper.synctex.gz | 13.175 | 12.835 | WinRAR archive | 17/01/2019 5:08 | 710A7B55 |
| abstract.tex | 90 | 76 | Archivo TEX | 17/01/2019 5:08 | 3C55448D |
| background.tex | 42 | 34 | Archivo TEX | 17/01/2019 5:08 | 0CD3476F |
| conclusion.tex | 43 | 35 | Archivo TEX | 17/01/2019 5:08 | 94B0B4EC |
| discussion.tex | 42 | 33 | Archivo TEX | 17/01/2019 5:08 | B3DF4B47 |
| IEEEtran.cls | 206.573 | 52.053 | Archivo CLS | 17/01/2019 5:08 | F53E2AA7 |
| introduction.tex | 46 | 33 | Archivo TEX | 17/01/2019 5:08 | 24EF0307 |
| paper.aux | 672 | 190 | Archivo AUX | 17/01/2019 5:08 | 2598A2A1 |
| paper.log | 51.533 | 7.134 | Documento de texto | 17/01/2019 5:08 | CC5EF946 |
| paper.pdf | 1.448.711 | 530.236 | Adobe Acrobat Do... | 17/01/2019 5:08 | 222FE8BB |
| paper.tex | 2.899 | 406 | Archivo TEX | 17/01/2019 5:08 | 80EC6402 |
| results.log | 444.486 | 10.130 | Documento de texto | 17/01/2019 5:08 | B63BB746 |
| results.tex | 11.541 | 1.570 | Archivo TEX | 17/01/2019 5:08 | 7C2B0B5C |
| threats.tex | 48 | 45 | Archivo TEX | 17/01/2019 5:08 | AD3BC8A4 |

Figura 34. Actividad respaldar resultados desde standalone

5.2.3. Análisis de resultados

Las encuestas fueron aplicadas a cinco investigadores pertenecientes a dos universidades. Cuatro de ellos fueron docentes de la Universidad de las Fuerzas Armadas ESPE, mientras que uno fue de la Universidad de Atacama de Chile. Para aplicar estas encuestas se acudió al sitio de trabajo de los investigadores locales y se les pidió que utilizaran la herramienta, posterior a ello respondieron la encuesta. Por otro lado, al investigador ubicado fuera del país se lo contactó por Skype y llenó la encuesta mediante un documento de texto. Las respuestas de los investigadores se analizan en las siguientes secciones de acuerdo a cada pregunta y todas las encuestas contestadas se encuentran en el Anexo 1.

5.2.3.1. Universidad a la que pertenece

Las respuestas a esta pregunta se encuentran en la Tabla 33, de la misma que se observa que el 80% de los encuestados son docentes de la Universidad de las Fuerzas Armadas ESPE, mientras que el 20% pertenece a la Universidad de Atacama, ubicada en Chile. Los resultados se los puede observar un gráfico tipo pastel en la Figura 35.

Tabla 33

Resultados universidad a la que pertenecen los encuestados

| CATEGORÍA | FRECUENCIA | PORCENTAJE |
|---|------------|------------|
| Universidad de las Fuerzas Armadas ESPE | 4 | 80% |
| Universidad de Atacama | 1 | 20% |

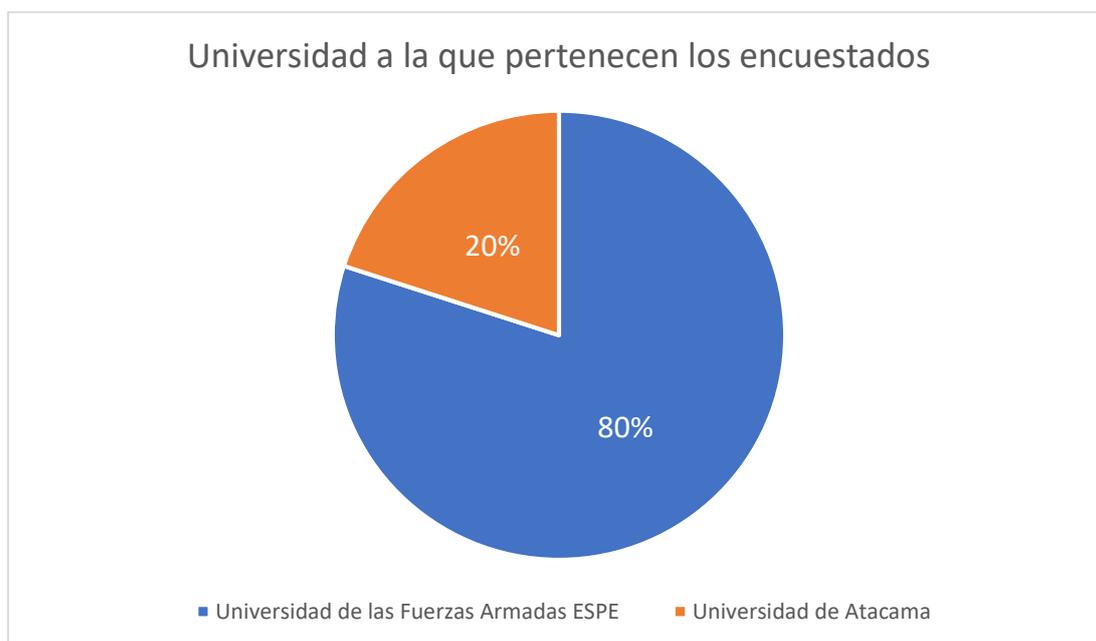


Figura 35. Universidad a la que pertenecen los encuestados

5.2.3.2. Pregunta 1

¿Considera usted que la herramienta propuesta facilita la aplicación del paradigma de investigación reproducible al caso de estudio “NaPiRE”? Justifique su respuesta.

Al ser una pregunta abierta, los resultados fueron tabulados en primera instancia, para luego ser presentados como se muestra en la Tabla 34.

Tabla 34

Resultados primera pregunta encuesta

| CATEGORÍA | FRECUENCIA | PORCENTAJE |
|--|------------|------------|
| Ayuda a replicar el estudio | 3 | 60% |
| Mantiene similitud entre replicaciones | 2 | 40% |

De la Tabla 34 se observa que tres de los cinco encuestados está de acuerdo en que la herramienta ayuda a replicar el caso de estudio, mientras que dos de ellos opinan que mantiene la similitud entre replicaciones. Es por esto que se concluye que la herramienta sí facilita la aplicación del paradigma de investigación reproducible al caso de estudio “NaPiRE”. Los

resultados se pueden observar en la Figura 36.

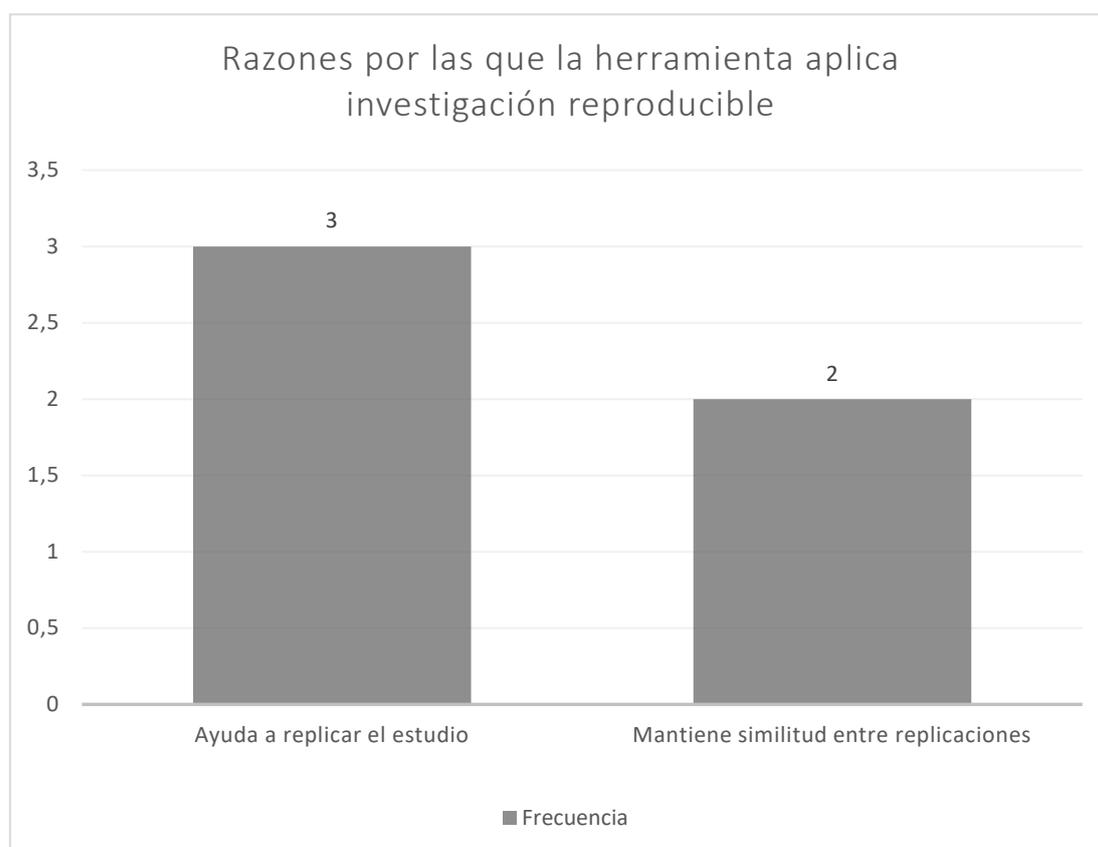


Figura 36. Razones por las que la herramienta apoya al caso de estudio

5.2.3.3. Pregunta 2

De las siguientes fases dentro de la investigación tradicional, encierre en un círculo el literal que considera que han sido agilizados gracias al prototipo de investigación reproducible. Considere como referencia la replicación de un estudio existente, mas no la creación de un nuevo estudio.

Los resultados de esta pregunta, luego de ser contabilizados, se encuentran en la Tabla 35.

Tabla 35
Resultados segunda pregunta encuesta

| RESPUESTA | FRECUENCIA | PORCENTAJE |
|------------|------------|------------|
| Resultados | 5 | 50% |

CONTINÚA →

| | | |
|-----------------------------------|---|-----|
| Observación, prueba y experimento | 3 | 30% |
| Definición del problema | 1 | 10% |
| Hipótesis | 1 | 10% |
| Razonamiento y deducción | 0 | 0% |

Después de analizar los resultados se encuentra que el total de los encuestados opinan que la herramienta agiliza la fase “Resultados”, tres de los cinco encuestados, opina que se agiliza la fase de “Observación, prueba y experimento”, tan solo uno en la fase de “Definición del problema” y de igual modo, uno en la fase de “Hipótesis”, dentro de la investigación tradicional. La Figura 37 muestra los resultados. De esto se concluye que la herramienta apoya fuertemente las fases de “Resultados” y “Observación, prueba y experimento”.

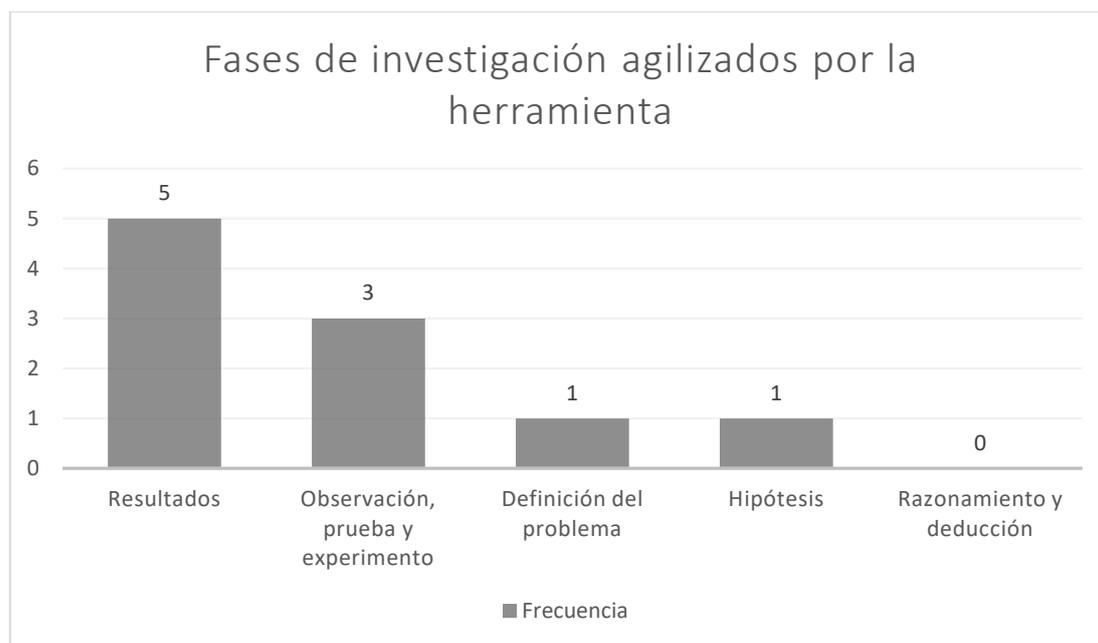


Figura 37. Fases de la investigación agilizados por la herramienta

5.2.3.4. Pregunta 3

¿Cómo le ayudó la herramienta propuesta a replicar el caso de estudio “NaPiRE”?

Como primer paso para analizar esta respuesta se categorizó las respuestas a esta pregunta abierta, después se elaboró la Tabla 36 con los resultados obtenidos.

Tabla 36
Resultados tercera pregunta encuesta

| CATEGORÍA | FRECUENCIA | PORCENTAJE |
|--|------------|------------|
| Automatización de análisis de datos y resultados | 3 | 60% |
| Optimización de tiempo empleado en el estudio | 2 | 40% |

Después de contabilizar los resultados, se observa que tres de cinco encuestados opinan que la herramienta ayuda en la “automatización de análisis de datos y resultados”, mientras que el resto opina que “optimiza el tiempo empleado en el estudio”. En base a estos resultados se concluye que la herramienta propuesta ayuda con la automatización de análisis de datos y resultados dentro de la replicación del caso de estudio “NaPiRE”. La Figura 38 ilustra los resultados a esta pregunta.

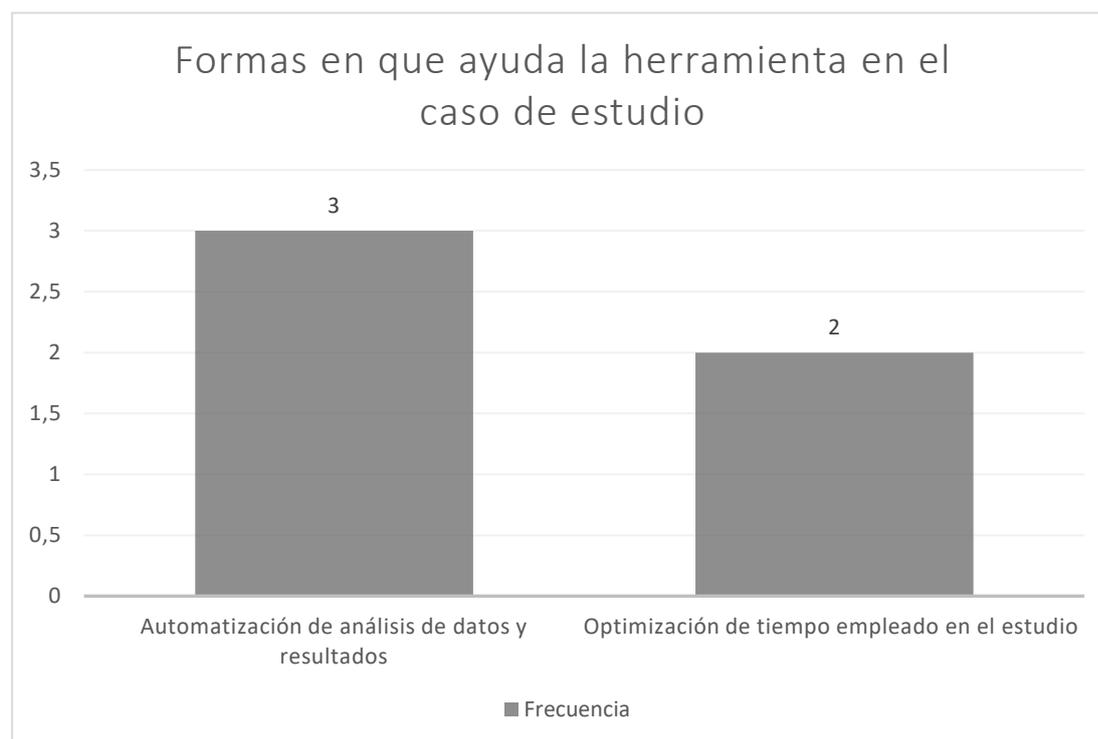


Figura 38. Formas en que ayuda la herramienta en el caso de estudio

5.2.3.5. Pregunta 4

Mediante la escala presentada, evalúe los beneficios de la investigación reproducible que encontró en la herramienta propuesta.

Las opciones de esta pregunta están basadas en la escala de Likert. Cada beneficio de la investigación reproducible es analizado independientemente.

a. La herramienta le ayuda a entender el proceso completo seguido durante la investigación.

Con los resultados se elaboró la Tabla 37.

Tabla 37
Resultados cuarta pregunta encuesta literal a

| RESPUESTA | CANTIDAD | PORCENTAJE |
|-------------------|----------|------------|
| De acuerdo | 3 | 60% |
| Muy de acuerdo | 2 | 40% |
| Indiferente | 0 | 0% |
| Desacuerdo | 0 | 0% |
| Muy en desacuerdo | 0 | 0% |

Después de procesar los datos se concluye que el 60% de los encuestados está *de acuerdo* en que la herramienta ayuda a entender el proceso completo que se llevó a cabo durante la investigación, mientras que un 40% indica estar *muy de acuerdo* con esta afirmación. De este análisis se concluye que la herramienta ayuda a entender el proceso completo seguido durante la investigación.

b. La herramienta proporciona los archivos utilizados en la investigación.

Los resultados de este literal se encuentran en la Tabla 38.

Tabla 38
Resultados cuarta pregunta encuesta literal b

| RESPUESTA | CANTIDAD | PORCENTAJE |
|-------------------|----------|------------|
| Muy de acuerdo | 4 | 80% |
| De acuerdo | 1 | 20% |
| Indiferente | 0 | 0% |
| Desacuerdo | 0 | 0% |
| Muy en desacuerdo | 0 | 0% |

Después de procesar los datos se concluye que el 80% de encuestados está *muy de acuerdo* en que la herramienta proporciona los archivos utilizados en la investigación, mientras que un 20% se muestra solamente *de acuerdo* con esta afirmación. De este análisis se concluye que la herramienta proporciona los archivos utilizados en la investigación.

c. La herramienta muestra el diseño de la encuesta utilizada en el estudio.

Las respuestas a este literal se encuentran en la Tabla 39.

Tabla 39
Resultados cuarta pregunta encuesta literal c

| RESPUESTA | CANTIDAD | PORCENTAJE |
|-------------------|----------|------------|
| Muy de acuerdo | 4 | 80% |
| De acuerdo | 1 | 20% |
| Indiferente | 0 | 0% |
| Desacuerdo | 0 | 0% |
| Muy en desacuerdo | 0 | 0% |

Después de procesar los datos se concluye que el 80% de encuestados está *muy de acuerdo* en que la herramienta muestra el diseño de la encuesta utilizada, mientras que un 20% se muestra solamente *de acuerdo* con esta afirmación. De este análisis se concluye que la herramienta muestra el diseño de la encuesta utilizada en el estudio.

d. La herramienta aumenta la confianza en los resultados del caso de estudio

Con los resultados se elaboró la Tabla 40.

Tabla 40

Resultados cuarta pregunta encuesta literal d

| RESPUESTA | CANTIDAD | PORCENTAJE |
|-------------------|----------|------------|
| De acuerdo | 3 | 60% |
| Muy de acuerdo | 2 | 40% |
| Indiferente | 0 | 0% |
| Desacuerdo | 0 | 0% |
| Muy en desacuerdo | 0 | 0% |

Después de procesar los datos se concluye que el 60% de encuestados está *de acuerdo* en que la herramienta aumenta la confianza en los resultados del caso de estudio, mientras que un 40% se muestra *muy de acuerdo* con esta afirmación. De este análisis se concluye que la herramienta aumenta la confianza en los resultados del caso de estudio.

e. La herramienta ayuda a contribuir con nuevos resultados apegados al marco común del caso de estudio principal

Los resultados de este literal se encuentran en la Tabla 41.

Tabla 41

Resultados cuarta pregunta encuesta literal e

| RESPUESTA | CANTIDAD | PORCENTAJE |
|-------------------|----------|------------|
| De acuerdo | 3 | 60% |
| Muy de acuerdo | 2 | 40% |
| Indiferente | 0 | 0% |
| Desacuerdo | 0 | 0% |
| Muy en desacuerdo | 0 | 0% |

Después de procesar los datos se concluye que el 60% de encuestados está *de acuerdo* en que la herramienta ayuda a contribuir con nuevos resultados apegados al marco común del

caso de estudio principal, mientras que un 40% se muestra *muy de acuerdo* con esta afirmación. De este análisis se concluye que la herramienta ayuda a contribuir con nuevos resultados apegados al marco común del caso de estudio principal.

En la Figura 39 se presentan todos los beneficios que ha brindado la herramienta compactados en un solo gráfico.

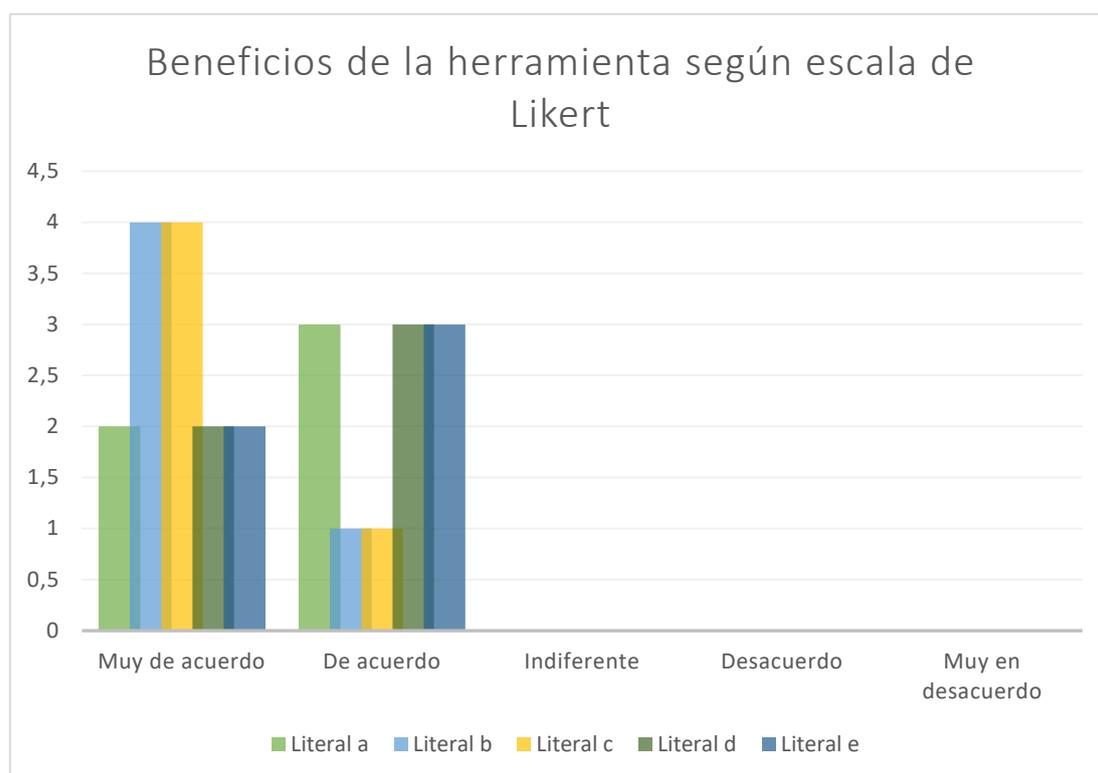


Figura 39. Beneficios de la herramienta según escala de Likert

CAPÍTULO VI

CONCLUSIONES, RECOMENDACIONES Y TRABAJOS FUTUROS

6.1 CONCLUSIONES

Existe una cantidad limitada de estudios relevantes en cuanto a la aplicación de la investigación reproducible en estudios empíricos de Ingeniería de Software. La mayoría de herramientas existentes utilizan línea de comandos para su ejecución por lo que su usabilidad se ve fuertemente afectada.

Se encontraron estudios que proponen algunos métodos y formas de replicación que han sido utilizadas, donde se identificó que la principal limitación es que no se ha podido establecer alguna herramienta para realizar reproducibilidad total aplicada a un área específica.

Es casi imposible crear una herramienta general para investigación reproducible debido a la complejidad y algunos componentes únicos que conforman una investigación, no obstante, según la naturaleza de la investigación es posible crear una herramienta genérica para un determinado contexto que incorpore investigación reproducible.

Los beneficios que aporta el prototipo desarrollado a la ingeniería de software son: facilitar el proceso de replicación de un estudio de tipo encuesta, automatizar el proceso de análisis de datos, agilizar la presentación del reporte de resultados y facilitar el proceso de aplicación de las encuestas.

Para que una investigación pueda ser reproducible necesita ser planeada con ese propósito desde su diseño, así como también incorporar las actividades claves dentro del proceso de ejecución, como son: buscar el mismo objetivo, utilizar similares procedimientos de análisis de datos y analizar el conjunto resultante de replications.

La investigación reproducible se puede llevar a cabo utilizando herramientas actuales de software, de forma que se mantenga continuidad en la herramienta y no quede obsoleta en

corto tiempo, a diferencia de algunas herramientas encontradas en la revisión sistemática de literatura.

6.2 RECOMENDACIONES

Se recomienda utilizar herramientas y componentes modernos al momento de desarrollar la herramienta ya que pueden existir problemas entre compatibilidades con OS servers por falta de complementos, librerías, etc.

Al momento de diseñar el estudio de tipo encuesta se recomienda aplicar estándares de modo que toda la documentación elaborada sea coherente y normalizado.

En la fase de diseño de un estudio es muy importante definir la forma correcta en que serán procesados los de datos de las encuestas según el tipo de pregunta, de esta forma se ayuda a que el estudio sea reproducible desde su diseño.

Dentro del repositorio interno que tiene la herramienta se recomienda manejar un registro de versiones dentro de los archivos del reporte de resultados de forma que se facilite el análisis individual o grupal de las replicaciones.

6.3 TRABAJOS FUTUROS

La herramienta propuesta deja abierta la posibilidad de incorporar la reproducibilidad de nuevos tipos de estudios, de esa forma se puede desarrollar un nuevo modelo que satisfaga las necesidades de otras clases de estudios.

Se puede mejorar el tiempo de respuesta de la herramienta propuesta optimizando la interacción con el servidor de R, es decir, perfeccionar la conexión entre Java y R o incluso desarrollar en otro lenguaje más eficiente.

Dentro de la herramienta se puede desarrollar o utilizar un editor de artículos científicos, de manera que la edición de los reportes de resultados sea en línea e integrado en la plataforma web.

A futuro en cuanto a la investigación reproducible se puede aplicar otra tecnología, por ejemplo: utilizar servidores de máquinas virtuales que contengan todo el estudio y las herramientas involucradas y que a partir de esta instancia se creen nuevas replicasiones.

REFERENCIAS BIBLIOGRÁFICAS

- Álvarez de Zayas, C., & Sierra, V. (1998). Como se modela la investigación científica. *La Investigación Científica En La Sociedad Del Conocimiento*, 1–155. Retrieved from [ecaths1.s3.amazonaws.com/tesis/48107023.UI_LIBRO BASE.pdf](http://ecaths1.s3.amazonaws.com/tesis/48107023.UI_LIBRO%20BASE.pdf)
- Cassidy, D., Holton, G., & Retherford, J. (2013). *Understanding Physics* (Springer S). New York.
- Cataldo, A. (2015). Design Science Research (DSR): Una breve introducción. *II Workshop RedSTI*, (February).
- Chairman, P. J. D., Comer, D. E., Gries, D., Mulder, M. C., Tucker, A., Turner, A. J. O. E., & Young, P. R. (1989). Computing as a Discipline. *Report of the ACM Task Force on the Core of Computer Science*, 32(1).
- Claerbout, J. (1991). Reproducible computational research: A history of hurdles, mostly overcome. *Stanford University*.
- da Silva, F. Q. B., Suassuna, M., França, A. C. C., Grubb, A. M., Gouveia, T. B., Monteiro, C. V. F., & dos Santos, I. E. (2012). *Replication of empirical studies in software engineering research: a systematic mapping study*. *Empirical Software Engineering*. <https://doi.org/10.1007/s10664-012-9227-7>
- Desikan, S., & Ramesh, G. (2008). *Software Testing: Principles and Practices* (Sixth Impr). Dorling Kindersley.
- Easterbrook, S., Singer, J., Storey, M.-A., & Damian, D. (2008). Selecting Empirical Methods for Software Engineering Research. *Guide to Advanced Empirical Software Engineering*, 285–311. https://doi.org/10.1007/978-1-84800-044-5_11
- Fernández, E. (2011). Comparative analysis of meta-analysis methods : when to use which?, 36–45.
- Fomel, S. (2015). Reproducible Research as a Community Effort: Lessons from the Madagascar Project. *SCIENTIFIC SOFTWARE COMMUNITIES*, (February). Retrieved from <http://ieeexplore.ieee.org/document/6880240/>

- Fomel, S., Sava, P., Vlad, I., Liu, Y., & Bashkardin, V. (2013). Madagascar: open-source software project for multidimensional data analysis and reproducible computational experiments. *Journal of Open Research Software*, 1(1), e8. <https://doi.org/http://dx.doi.org/10.5334/jors.ag>
- Fonseca Carrera, E. R. (2012). Definition of a support infrastructure for replicating and aggregating families of software engineering experiments. In *International Doctoral Symposium on empirical Software Engineering* (p. 8). Lund.
- Gauch, H. G. (2003). *Scientific Method in Practice*. (C. U. Press, Ed.) (2003rd ed.). The press syndicate of the University of Cambridge.
- Goddard, W., & Melville, S. (2001). *Research Methodology: An Introduction*.
- Hernández León, R. A., & Coello Gonzáles, S. (2011). El Proceso de Investigación Científica.
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2010). *Metodología de la Investigación* (Quinta edi). México: The McGraw-Hill.
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2014). Metodología de la investigación. *Journal of Chemical Information and Modeling*, 53(9), 1–589. <https://doi.org/10.1017/CBO9781107415324.004>
- Hevner, A. R., March, S. T., Park, J. & Ram, S. (2004). Design science in information systems research, (September 2014), 75–105.
- Iqbal, S. A., Wallach, J. D., Khoury, M. J., & Ioannidis, J. P. (2016). Reproducible Research Practices and Transparency across the Biomedical Literature. *PLoS Bios*.
- Kitchenham, B., & Charters, S. (2007). *Guidelines for performing Systematic Literature reviews in Software Engineering Version 2.3. Engineering* (Vol. 45). <https://doi.org/10.1145/1134285.1134500>
- Kitzes, J., Turek, D., & Deniz, F. (Eds.). (2018). *The Practice of Reproducible Research*. University of California Press.
- Kothari, C. R. (2004). *Research Methodology: Methods & Techniques*. New Age International (P) Ltd. <https://doi.org/10.1017/CBO9781107415324.004>

- Larsson, H., Lindqvist, E., & Torkar, R. (2014). Outliers and replication in software engineering. *Proceedings - Asia-Pacific Software Engineering Conference, APSEC, 1*, 207–214. <https://doi.org/10.1109/APSEC.2014.40>
- Linares Fontela, J. (2015). *Guía para diseñar encuestas*.
- Madrid, J. P., & Macchetto, D. (2009). High-Impact Astronomical Observatories. <https://doi.org/10.1126/science.1213443>
- Martel, A. (2014). *Gestión Práctica de Proyectos Con Scrum: Desarrollo de Software Ágil Para El Scrum Master*. CreateSpace Independent Publishing Platform.
- Martinez, R., & Fernández, A. (2015). Metodologías e instrumentos para la formulación, evaluación y monitoreo de programas sociales, 2–6. Retrieved from http://recursos.salonesvirtuales.com/assets/bloques/martinez_rodrigo.pdf
- Méndez, D., Wagner, S., Felderer, M., & Kalinowski, M. (2012). Manifesto.
- Méndez Fernández, D., Wagner, S., Kalinowski, M., R, J. C. O., Conte, T., & Prikladmicki, R. (2015). Towards Building Knowledge on Causes of Critical Requirements Engineering Problems. *International Conference on Software Engineering and Knowledge Engineering*. [https://doi.org/DOI: 10.18293/SEKE2015-220](https://doi.org/DOI:10.18293/SEKE2015-220)
- Meng, H., Thain, D., Vyushkov, A., Wolf, M., & Woodard, A. (2016). Conducting Reproducible Research with Umbrella : Tracking , Creating , and Preserving Execution Environments. *2016 IEEE 12th International Conference on EScience*, 91–100. <https://doi.org/10.1109/eScience.2016.7870889>
- Meng, H., Thain, D., Vyushkov, A., Wolf, M., & Woodard, A. (2017). Conducting reproducible research with Umbrella: Tracking, creating, and preserving execution environments. *Proceedings of the 2016 IEEE 12th International Conference on E-Science, e-Science 2016*, 91–100. <https://doi.org/10.1109/eScience.2016.7870889>
- Monperrus, M. (2016). Introduction to Empirical Software Engineering 2 A Tentative Definition of Empirical Software Engineering. *University of Lille*, 1–13.
- Morgan, M. (2014). Reproducible Research. *Computing in Science & Engineering*, 12(5),

2014–2015. <https://doi.org/10.1109/MCSE.2010.113>

Pearson, K. (1999). *The Grammar of Science*, 1–16.

Peng, R. D. (2011). Reproducible Research in Computational Science. *Biostatistics*, *10*(3), 405–408. <https://doi.org/10.1126/science.1213847>.Reproducible

Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2008). Systematic Mapping Studies in Software Engineering. *12Th International Conference on Evaluation and Assessment in Software Engineering*, *17*, 10. <https://doi.org/10.1142/S0218194007003112>

Pressman, R. S. (2010). *Ingeniería del software. Un enfoque práctico*. <https://doi.org/http://zeus.inf.ucv.cl/~bcrawford/Modelado%20UML/Ingenieria%20del%20Software%207ma.%20Ed.%20-%20Ian%20Sommerville.pdf>

Sáenz Campos, D., & Tinoco Mora, Z. (1999). Introducción a la Investigación Científica, *12*, 60–71.

Schmidt, S. (2009). Shall We Really Do It Again? The Powerful Concept of Replication Is Neglected in the Social Sciences. *Review of General Psychology*, *13*(2), 90–100. <https://doi.org/10.1037/a0015108>

Schulte, E., Davison, D., Dye, T., & Dominik, C. (2012). A Multi-Language Computing Environment for Literate Programming and Reproducible Research. *J Stat Software*, *46*(3), 1–24. <https://doi.org/10.18637/jss.v046.i03>

Shields, P. M., & Rangarajan, N. (2013). *A Playbook for Research Methods: Integrating Conceptual Frameworks and Project Management*. New Forums Press. Retrieved from https://www.researchgate.net/publication/263046108_A_Playbook_for_Research_Methods_Integrating_Conceptual_Frameworks_and_Project_Management

Sjøberg, D. I. K., Hannay, J. E., Hansen, O., Kampenes, V. B., Karahasanovic, a. Liborg, N. K., & Rekdal, a. C. (2005). A Survey of 116 Controlled Experiments in Software Engineering. *IEEE Transactions on Software Engineering*, *31*(9), 733–753.

Sommerville, I. (2011). *Software Engineering*. Pearson.

Sutherland, J. (2015). *Scrum: The Art of Doing Twice the Work in Half the Time*. Scrum Inc.

- Undersea, N., Northwest, P., & Bauer, S. (2013). What is science? Discovery : The spark for science. *University of California Museum of Paleontology*, 2014, 1–22.
- Van Aken, J. E. (2005). Management research as a design science: Articulating the research products of mode 2 knowledge production in management. *British Journal of Management*, 16(1), 19–36. <https://doi.org/10.1111/j.1467-8551.2005.00437.x>
- Wieringa, R. (2013). Introduction to design science methodology. *REFSQ Doctoral Symposi*, 1–17.
- Wieringa, R. (2014). *Design Science Methodology for Information Systems and Software Engineering. Proceedings of the 32nd ACM/IEEE International Conference on Software Engineering - ICSE '10* (Vol. 2). <https://doi.org/10.1145/1810295.1810446>
- Zapata Domínguez, Á., Murillo Vargas, G., & Martínez Crespo, J. (2009). Teorías Contemporáneas de la Organización y del Management.
- Zendler, A. (2001). A preliminary software engineering theory as investigated by published experiments. *Empirical Software Engineering*, 6(2), 161–180. <https://doi.org/10.1023/A:1011489321999>