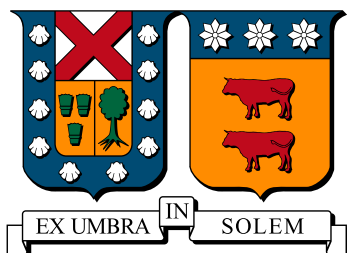


UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INFORMÁTICA
VALPARAÍSO – CHILE



**“DISEÑO DE UN PROTOCOLO
ESTANDARIZADO DE EVALUACIÓN DE
EXPERIENCIA DE USUARIO PARA
PROYECTOS DESARROLLADOS PARA LA
FERIA DE SOFTWARE.”**

CAMILA ANTONIA MENA ARAYA

**MEMORIA DE TITULACIÓN PARA OPTAR AL TÍTULO DE
INGENIERA CIVIL INFORMÁTICA**

PROFESOR GUÍA: PEDRO GODOY BARRERA

PROFESOR CORREFERENTE: WLADIMIR ORMAZÁBAL ORELLANA

OCTUBRE 2026

MATERIAL DE REFERENCIA, SU USO NO INVOLUCRA RESPONSABILIDAD DEL AUTOR O DE LA INSTITUCIÓN



CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): ☒ Memoria o trabajo de título ☐ Tesis de Postgrado

Título del trabajo: Diseño de un protocolo estandarizado de evaluación de Experiencia de Usuario para proyectos desarrollados para la Feria de Software

Nombre del candidato(a): Camila Antonia Mena Araya

Carrera / Grado: Ingeniería Civil Informática

Campus: Casa Central **Departamento:** Informática

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, Pedro Francisco Godoy Barrera, en mi calidad de profesor(a) guía/director(a) del trabajo académico mencionado anteriormente **DEJO CONSTANCIA** que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución.

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL (marcar una opción)

☒ El trabajo **NO contiene** información que amerite confidencialidad y puede ser publicado de inmediato en repositorio con acceso abierto.

☐ El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad (**embargo**) por (**marcar una opción**):

☐ 6 meses ☐ 12 meses ☐ 2 años ☐ 3 años ☐ 5 años ☐ 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

4.- FIRMAS

Profesor(a) guía o director(a) de memoria o tesis:

Fecha: 07-Octubre-2026

Firma:

Estudiante o Candidato(a):

Fecha: 07/10/2026

Firma:

Este formulario debe ser insertado como página 2 de la memoria o tesis, completado y firmado por estudiante y profesor(a) antes de la entrega en portal PRISMA de Biblioteca USM.

Agradecimientos

Dedico esta memoria, en primer lugar, a mi abuelo. Aunque no alcanzó a verme entrar a la universidad, desde mis días en el preuniversitario tuvo siempre la inquebrantable certeza de que lograría llegar hasta aquí.

A mi mamá, por su esfuerzo incondicional para que nunca me faltara nada. Verla levantar su negocio desde cero y tener la oportunidad de ayudarla me entregó conocimientos y fortalezas invaluable. Ella me enseñó que, sin importar las adversidades, somos fuertes y podemos salir adelante si le ponemos dedicación y cariño a lo que hacemos.

A mi abuela, que siempre estuvo ahí con un almuerzo rápido para que pudiera seguir con mis clases, o con uno de los postres que tanto me gustan, brindándome energía y cariño. A toda mi familia, por ser mi red de apoyo constante.

A mi pololo, Joshe, por ser un pilar fundamental en mi avance por la carrera. Gracias por estar a mi lado y aconsejarme cada vez que sentí que un ramo me superaba. Incluso en mis días más pesimistas, siempre lograste mostrarme el lado bueno de las cosas, ayudándome a recuperar la calma cuando los pensamientos me ganaban y, sobre todo, por ayudarme a ser una mejor persona. Extiendo este agradecimiento a su grupo de amigos, quienes me han acogido con los brazos abiertos y me han hecho sentir como una más en cada uno de nuestros planes.

A mis profesores de la Escuela de Gimnasia Artística Valparaíso, quienes, sin saberlo, se convirtieron en un soporte inmenso frente al estrés universitario. Gracias por las conversaciones durante los entrenamientos, por sacarme miles de risas, por permitirme desconectar cuando más lo necesitaba y por ayudarme a cumplir el gran sueño de poder competir.

A mis amigos que me apañaron tantas horas en Discord, ya sea jugando o viéndolos jugar. Gracias por las risas, por estar presentes y por hacerme sentir parte del grupo, a pesar de no haber comenzado la carrera junto a ustedes.

A mi segunda familia, los de la JIM/BIS, quienes fueron una parte esencial de mi crecimiento en la universidad. Con ustedes pude ser yo misma y disfrutar plenamente esta etapa. Sin importar el momento ni el lugar, siempre los llevaré en mi corazón y me alegrará verlos.

A todos los profesores y profesoras de la universidad que no dudaron en tenderme una mano cuando más lo necesité. Agradezco profundamente su disposición, paciencia y apoyo constante, los cuales fueron fundamentales para ayudarme a salir adelante y superar con éxito las asignaturas que más me costaron a lo largo de la carrera.

A las queridas mascotas que me acompañaron en este trayecto: a mi perrita Mini, que estuvo a mi lado durante los duros primeros años de estudio; a mi perrito Caramelo, que llegó para llenar un vacío y acompañarme en las largas noches de desvelo; y, finalmente, al gato negro de mi edificio, al que dejaba entrar en mis noches de mayor estrés para encontrar consuelo y compañía en sus ronroneos.

Por último, quiero dedicar esta memoria a mí misma. Muchas veces llegué a dudar de que este momento llegaría, por lo que quiero dejar estas palabras como un recordatorio personal: siempre se puede. Todos tenemos fortalezas en nuestro interior; la clave está en poner esfuerzo, cariño y dedicación.

Resumen

En la “*Feria de Software de la Universidad Técnica Federico Santa María*”, los estudiantes desarrollan proyectos informáticos de alta complejidad funcional orientados a la innovación. Sin embargo, la Experiencia de Usuario suele ser postergada debido a la urgencia de la implementación técnica y a la gran heterogeneidad de los ecosistemas desarrollados (aplicaciones web, móviles, inmersivas, IoT, sistemas críticos, entre otros), lo que hace ineficaz el uso de métricas de usabilidad genéricas.

Para resolver esta brecha metodológica, el presente trabajo propone el diseño de un protocolo de evaluación UX estandarizado y adaptable. La metodología incluyó una revisión del estado del arte y la categorización empírica de proyectos, consolidando una taxonomía de ocho ecosistemas tecnológicos. A cada uno se le asignó estratégicamente una dupla de métricas psicométricas validadas mediante un modelo de triangulación (micro-táctica y macro-estratégica). Esta arquitectura permite aislar y medir con precisión factores críticos de frontera, tales como la carga cognitiva, la fatiga ergonómica o la confianza algorítmica, ajustándose a la naturaleza del software.

El protocolo guía operativamente a los equipos mediante un Árbol de Decisión Diagnóstico y estandariza el reporte de resultados a través de un *Framework* denominado *UX-Scorecard*. La validación del instrumento se ejecutó bajo el Modelo de Aceptación Tecnológica (TAM) en la cohorte 2026, complementada con una Prueba de Concepto mediante estudios de caso iterativos. Los resultados demostraron una alta utilidad diagnóstica y facilidad de adopción, concluyendo que esta herramienta transforma el control de calidad en las asignaturas, permitiendo a los estudiantes fundamentar sus decisiones de rediseño en evidencia empírica rigurosa.

Abstract

At the “*Feria de Software de la Universidad Técnica Federico Santa María*”, students develop functionally complex computing projects aimed at innovation. However, User Experience is often deprioritized due to the urgency of technical implementation and the wide heterogeneity of the developed ecosystems (web, mobile, immersive, IoT, and critical systems applications, among others), which makes the use of generic usability metrics ineffective.

To address this methodological gap, this work proposes the design of a standardized and adaptable UX evaluation protocol. The methodology included a review of the state of the art and the empirical categorization of projects, consolidating a taxonomy of eight technological ecosystems. Each ecosystem was strategically assigned a pair of validated psychometric metrics through a triangulation model (micro-tactical and macro-strategic). This architecture makes it possible to precisely isolate and measure critical boundary factors, such as cognitive load, ergonomic fatigue, or algorithmic trust, tailored to the nature of the software.

The protocol operationally guides teams through a Diagnostic Decision Tree and standardizes results reporting through a framework called the UX-Scorecard. The instrument was validated under the Technology Acceptance Model (TAM) with the 2026 cohort, complemented by a Proof of Concept using iterative case studies. The results demonstrated high diagnostic utility and ease of adoption, concluding that this tool transforms quality control in coursework, enabling students to base their redesign decisions on rigorous empirical evidence.

Índice

Agradecimientos	III
Resumen	V
Abstract	VI
Índice	VII
Introducción	1
1. Definición del Problema	3
1.1. Objetivos Generales	9
1.2. Objetivos Específicos	9
2. Estado del arte	10
2.1. Análisis Comparativo de Principios de Usabilidad	11
2.1.1. Selección y Fundamentación del Marco Heurístico	13
2.2. Paradigmas de Adaptabilidad y Dependencia Tecnológica en la Evaluación de UX	14
2.2.1. Determinismo del Soporte Tecnológico y el Contexto de Uso en la Interfaz	14

2.2.2.	Marcos de Trabajo Flexibles y Estandarización de la Calidad del Software	17
2.3.	Paradigma de Triangulación y Enfoque Plural en la Medición de UX	19
2.3.1.	El Paradigma de Denzin: Principios de la Triangulación Científica .	19
2.4.	Escalas de evaluación para la medición de la UX según su entorno tecnológico	21
2.4.1.	Aplicaciones Web	22
2.4.2.	Aplicaciones móviles	22
2.4.3.	Videojuegos	23
2.4.4.	Realidad Virtual y Simulación	24
2.4.5.	Dispositivos Vestibles (Wearables)	26
2.4.6.	Sistemas Automatizados e Inteligencia Artificial	26
2.4.7.	Métricas Transversales y Marcos Complementarios de UX	27
2.5.	Consolidación de instrumentos de medición según entorno tecnológico . . .	31
3.	Propuesta de Solución	34
3.1.	Paso 1: Conceptualización y categorización preliminar	35
3.1.1.	Análisis de proyectos presentados en la Feria de Software 2024 y tecnología predominante de cada proyecto:	36
3.2.	Paso 2: Validación de la categorización preliminar	41
3.2.1.	Análisis de Hallazgos de la Encuesta	44
3.2.2.	Refinamiento de Ecosistemas Existentes	46
3.3.	Paso 3: Reestructuración de las categorías	48
3.4.	Paso 4: Asignación estratégica de metodologías y métricas de UX	50
3.4.1.	Criterios de Selección y Modelo de Triangulación Metodológica . .	51
3.4.2.	Clasificación de Ecosistemas Tecnológicos y Matriz de Escalas UX	53
3.5.	Paso 5: Diseño del protocolo de evaluación UX para proyectos de la FESW	63

3.5.1.	Paso 5: Diseño del protocolo de evaluación UX para proyectos de la FESW	63
3.5.2.	Ruta Diagnóstica: Árbol de Decisión para la Asignación de Ecosistemas	65
3.5.3.	Framework de Evaluación: UX-Scorecard de Trazabilidad	69
4.	Validación	72
4.1.	Caracterización de la Muestra	73
4.2.	Resultados Cuantitativos: Modelo de Aceptación Tecnológica	74
4.2.1.	Análisis Descriptivo de la Aceptación Tecnológica	75
4.2.2.	Análisis Cualitativo y Oportunidades de Mejora	76
4.2.3.	Validación Final: Prueba de Concepto (Estudios de Caso)	78
4.2.4.	Síntesis y Cierre de la Validación Empírica	82
	Conclusiones y trabajo futuro	83
5.1.	Conclusiones Generales	83
5.1.1.	Cumplimiento de los Objetivos de la Investigación	83
5.1.2.	Reflexiones Metodológicas y Prácticas	85
5.2.	Trabajo Futuro	86
	Bibliografía	88
A.	Protocolo y Guías de Uso de las Escalas de Evaluación UX	91

Introducción

La formación en Ingeniería Civil Informática en la UTFSM encuentra su hito culminante en la Feria de Software, una instancia académica donde los estudiantes conciben, diseñan y desarrollan proyectos tecnológicos orientados al emprendimiento, la innovación y a la resolución de problemáticas reales. En el actual ecosistema digital, el éxito y la adopción de estas soluciones no dependen exclusivamente de su robustez técnica o del cumplimiento de requerimientos funcionales, sino que están intrínsecamente ligados a la calidad de la Experiencia de Usuario.

A pesar de la relevancia estratégica de esta disciplina, en el ciclo de desarrollo de las asignaturas asociadas a la FESW se evidencia una postergación sistemática de las evaluaciones de usabilidad frente a la urgencia de la implementación del código. Esta brecha metodológica se ve agudizada por la alta heterogeneidad de los proyectos presentados, los cuales abarcan desde plataformas web y aplicaciones móviles tradicionales, hasta entornos inmersivos de Realidad Virtual, dispositivos vestibles (*wearables*) y sistemas críticos impulsados por Inteligencia Artificial. Ante esta diversidad, la aplicación empírica de métricas de usabilidad genéricas o universales resulta ineficaz, lo que impide a los equipos auditar correctamente las restricciones físicas y cognitivas de sus propios desarrollos.

Para dar respuesta a esta problemática, la presente memoria de título propone el diseño de un protocolo de evaluación UX estandarizado, trazable y adaptable a múltiples ecosistemas tecnológicos. A través de la definición de una taxonomía estructurada y la asignación de duplas de escalas psicométricas validadas mediante triangulación metodológica (micro-táctica y macro-estratégica), se busca dotar a los estudiantes de un marco operativo concreto sistematizado en el *Framework UX-Scorecard* que les permita fundamentar sus decisiones de

rediseño en evidencia empírica rigurosa.

Para exponer el desarrollo de esta investigación, el documento se estructura en cuatro capítulos más las conclusiones y el trabajo futuro: El **Capítulo 1** define el problema, el contexto académico y los objetivos generales y específicos que guían el trabajo. El **Capítulo 2** presenta el estado del arte, revisando los marcos heurísticos y los instrumentos psicométricos de evaluación UX según su entorno tecnológico. El **Capítulo 3** detalla la propuesta de solución, abarcando desde la categorización empírica de los ecosistemas hasta la conformación del protocolo y la ruta diagnóstica. El **Capítulo 4** expone la validación del instrumento mediante el Modelo de Aceptación Tecnológica (TAM) y pruebas de concepto aplicadas a estudios de caso reales de la cohorte 2026. Finalmente, se presentan las **Conclusiones Generales 5.1** de la investigación, el grado de cumplimiento de los objetivos y las proyecciones de **Trabajo Futuro 5.2**.

Capítulo 1

Definición del Problema

La Feria de Software de ahora en adelante FESW, de la Universidad Técnica Federico Santa María es un hito tradicional enfocado en el “aprender haciendo”, el cual opera como el eje articulador del último año de la carrera de Ingeniería Civil Informática. Dicha articulación se manifiesta en una triple dimensión formativa: en primer lugar, integra y sintetiza transversalmente las competencias teóricas, técnicas y de gestión adquiridas por los estudiantes a lo largo de su plan de estudios; en segundo lugar, establece una continuidad metodológica y secuencial entre la formulación estratégica de soluciones y la ejecución técnica de software; y en tercer lugar, conecta el quehacer universitario con problemáticas reales del entorno productivo y social. La preparación formal de este evento comienza durante el primer semestre con la asignatura *INF-360: Gestión de Proyectos de Informática*. En esta etapa inicial, conocida también como el pre-proyecto, los estudiantes forman equipos con el objetivo de buscar problemáticas reales provenientes de clientes del sector público, privado o social; en este punto, resulta fundamental que las propuestas de software y el alcance tecnológico diseñado vayan estrechamente vinculados a los Objetivos de Desarrollo Sostenible (ODS), validando luego la viabilidad técnica y comercial ante una comisión docente y ayudantes de la asignatura. El objetivo principal de este período es construir y testear un Producto Mínimo Viable (MVP) completamente funcional y libre de maquetas. Posteriormente, durante el segundo semestre, el proceso de formación continúa formalmente con la asignatura *INF-228: Taller de Desarrollo de Proyectos de Informática*. En este ramo práctico, los alumnos

transforman sus equipos en “startups” de tecnología y entran en una fase de ejecución pura mediante metodologías ágiles de desarrollo. El objetivo principal de este período es pulir la tanto el back-end como la interface, gestionando de forma directa la relación con el cliente y enfrentando las exigencias de plazos propias del entorno profesional real.

Toda esta trayectoria académica confluye en el evento masivo de la Feria de Software, donde se juntan los alumnos de campus Casa Central y campus San Joaquín en una jornada abarca todo el día en la universidad a mediados de noviembre para abrir uno de los dos campus involucrados a la comunidad, colegios y empresas. En este espacio, los estudiantes montan stands y realizan presentaciones de tipo pitch para demostrar el funcionamiento de sus plataformas tecnológicas a todo el público donde entre ellos se encuentran comisiones evaluadoras compuestas por académicos y expertos de la industria. La experiencia cierra con instancias de vinculación como el Demo-Day y una premiación final, consolidándose como un rito de iniciación clave que los dejará a un paso de que los alumnos dejen atrás su rol de estudiante para comenzar a operar formalmente como ingenieros.

Sin embargo, para que estos productos concebidos a lo largo de este ciclo asignatural alcancen los estándares exigidos por el mercado los cuales contemplan la calidad de software y las buenas prácticas de diseño de la industria, y puedan destacar ante las comisiones del evento, no basta con cumplir únicamente con requerimientos funcionales o de programación. Es en la transición desde el desarrollo de código hacia la interacción con los usuarios finales donde emerge una brecha metodológica compleja dentro del proceso de formación de las asignaturas mencionadas. Para subsanar esta falencia, se vuelve indispensable incorporar formalmente el concepto de Experiencia de Usuario (UX de ahora en adelante), definida por el estándar ISO 9241-210:2019 como ***“las percepciones y respuestas de la persona resultantes del uso y/o del uso anticipado de un producto, sistema o servicio, abarcando todas las emociones, creencias, preferencias y respuestas físicas y psicológicas que surgen antes, durante y después de dicha interacción”***[1]. Al adentrarse en la literatura especializada, es evidente que coexisten múltiples aproximaciones teóricas y normativas sobre este término; algunas orientadas estrictamente a los aspectos cognitivos, otras enfocadas en los atributos hedónicos o estéticos del software, e incluso visiones estandarizadas a nivel internacional como las normas ISO. No obstante, dada la naturaleza de las asignaturas y el enfoque social y

en el emprendimiento de la feria, para efectos de este trabajo se adoptará de manera explícita la definición fundamental acuñada por los pioneros de la industria, Jakob Nielsen y Donald Norman. Según estos autores, ***“la experiencia de usuario abarca todos los aspectos de la interacción del usuario final con la empresa, sus servicios y sus productos”***[2]. La elección de esta postura teórica no es fortuita; radica en su carácter holístico y sistémico. Al establecer que la UX trasciende el diseño de una interfaz visual y engloba la totalidad del contacto entre la persona y la solución propuesta, esta definición proporciona el paraguas conceptual idóneo para analizar proyectos de software. De este modo, se asume que la experiencia no es un atributo aislado del código, sino el resultado integral de cómo el producto final responde a las necesidades reales del usuario, interactúa con su entorno y cumple con las promesas de valor técnico y operativo definidas desde las primeras etapas del pre-proyecto.

En los últimos años, la Experiencia de Usuario ha dejado de ser un complemento estético secundario para convertirse en un pilar estratégico y protagónico dentro del ciclo de desarrollo de software. Este giro responde a una transformación del mercado: en un ecosistema saturado de soluciones computacionales equivalentes, el éxito de un producto no depende únicamente de su robustez técnica o cumplimiento de requerimientos funcionales, sino de su capacidad de adopción y retención por parte del usuario final. Diversos estudios de la industria respaldan esta transición; investigaciones desarrolladas por McKinsey & Company evidencian que las organizaciones que integran rigurosamente el diseño centrado en el usuario superan financieramente a sus pares de la industria en un 32 % de crecimiento de ingresos [3]. Asimismo, desde la perspectiva de la ingeniería de software, la incorporación temprana de metodologías UX minimiza sustancialmente los costos de desarrollo y la deuda técnica, mitigando el riesgo de iteraciones tardías en etapas de producción donde el costo de modificación es exponencialmente más elevado [4].

Bajo este escenario, la formulación y gestión de proyectos informáticos innovadores, destinados a presentarse en instancias como la Feria de Software de la UTFSM, enfrentan un desafío crítico en la actualidad. Si bien estos desarrollos están fuertemente orientados al emprendimiento y a la resolución de problemáticas reales inspiradas por los Objetivos de Desarrollo Sostenible (ODS), su éxito, la adopción por parte de los usuarios y el verdadero impacto social en el entorno dependen intrínsecamente de una experiencia de usuario satisfactoria.

A pesar de la relevancia de esta disciplina, las asignaturas que concluyen en la Feria de

Software operan como cursos de tipo *capstone* es decir “*instancias académicas culminantes diseñadas para que los estudiantes integren, consoliden y apliquen de forma práctica las competencias adquiridas a lo largo de toda su carrera en un proyecto de carácter profesional*” [5]. Sin embargo, en el transcurso de su formación, los alumnos cursan únicamente una asignatura obligatoria vinculada a esta área: **INF-322 Diseño de Interfaces Usuaris.** Dicho ramo se centra principalmente en la capa visual y de interacción directa, careciendo de módulos mandatorios enfocados integralmente en la Experiencia de Usuario. Si bien la malla curricular ofrece la asignatura electiva “Métodos de Diseño Centrado en Usuarios”, su naturaleza opcional determina que una gran parte del alumnado no transite por ella, conformando una brecha de conocimiento sustancial en el área. Como resultado de esta falencia formativa, los estudiantes suelen carecer de conocimientos profundos respecto a las metodologías y evaluaciones de UX aplicables a sus desarrollos. Esto provoca que ciertos atributos de calidad esenciales, como la faceta de usabilidad de cara al usuario final, sean postergados o dejados de lado frente a la urgencia de la implementación técnica. Asimismo, los alumnos frecuentemente no tienen claridad sobre cómo ejecutar las técnicas de revisión de UX o desconocen los mecanismos para integrarlas en sus flujos de trabajo.

En el contexto de las asignaturas orientadas a la creación de proyectos, se observa que el ecosistema de las soluciones desarrolladas es altamente heterogéneo. Esta diversidad no solo se manifiesta en el soporte físico utilizado abarcando desde aplicaciones Web y Mobile tradicionales, hasta experiencias inmersivas de Realidad Virtual (VR) y dispositivos vestibles, sino también en la complejidad de sus arquitecturas lógicas y contextos operativos, como ocurre en los sistemas de Ciberseguridad o el análisis predictivo. Esta multidimensionalidad tecnológica vuelve ineficaz la aplicación de reglas uniformes para la evaluación de la UX, dado que cada proyecto presenta exigencias de interacción, cargas cognitivas y limitantes de hardware sustancialmente distintas entre sí.

Un ejemplo de esto es que, mientras que una aplicación web o móvil se mide por su eficiencia en completar tareas, un videojuego requiere evaluar el compromiso y el flujo para asegurar que sea divertido, y una experiencia de VR/AR debe priorizar la inmersión y la ergonomía cognitiva para evitar malestares físicos como el mareo por movimiento. En este último caso, un fallo en la UX no solo dificulta el uso, sino que puede traer consecuencias reales físicas al usuario, convirtiendo la performance en un aspecto crítico de la experiencia.

Actualmente, esta falta de estandarización y el desconocimiento metodológico se traducen en que muchos equipos carezcan de un proceso formal que conecte las decisiones de diseño con requerimientos específicos y pruebas posteriores, limitando la trazabilidad de la calidad del producto final. No se aplican heurísticas de evaluación adaptadas a la tecnología específica (como la escala de evaluación SUS para web, de Playability para juegos, o específicas para VR). Además, a menudo se omiten pruebas con usuarios reales o se carece de métricas objetivas para validar la efectividad de la UX.

Esta situación conlleva el riesgo de desarrollar productos con una usabilidad deficiente, lo que impacta negativamente en la adopción por parte de los usuarios y, en última instancia, en el éxito del proyecto en el mercado y en el cumplimiento efectivo de la solución planteada para la problemática ODS. La falta de un protocolo claro también dificulta la capacidad de los estudiantes para documentar y certificar la calidad de su trabajo de manera formal.

A partir de esta problemática surge una oportunidad clave: aprovechando la gran variedad de proyectos desarrollados en las asignaturas y la diversidad de herramientas de evaluación de UX existentes en la actualidad, es posible estructurar una solución sistemática. Esto evidencia la necesidad urgente de establecer un marco de trabajo común y un protocolo de trazabilidad para las pruebas de UX durante el desarrollo del curso. Dicho protocolo debe adaptarse a las distintas tecnologías utilizadas por los estudiantes; esto incluye realizar una clasificación exhaustiva de los proyectos según su naturaleza tecnológica para, posteriormente, proponer y asignar una escala de evaluación adecuada y específica para cada tipo. De este modo, se guiará el desarrollo mediante la definición de un framework de evaluación (como un UX-Scorecard) y la planificación de etapas claras para el plan de trazabilidad desde el inicio de la idea hasta el despliegue final.

A modo de síntesis, el análisis del contexto de desarrollo de proyectos para la Feria de Software evidencia una brecha metodológica crítica en el aseguramiento de la calidad desde la perspectiva del usuario final. Los principales problemas detectados se resumen en los siguientes puntos:

- **Carencia formativa transversal en Experiencia de Usuario:** Existe una postergación sistemática de los atributos de usabilidad frente a la urgencia de la implementación técnica, originada por la falta de instrucción obligatoria enfocada integralmente en metodologías de evaluación UX dentro de las asignaturas.
- **Heterogeneidad tecnológica sin métricas adaptadas:** La diversidad de ecosistemas de software desarrollados (Web, Mobile, Realidad Virtual, Videojuegos, entre otros) vuelve ineficaz la aplicación de reglas uniformes, dejando un vacío operativo sobre cómo medir de manera precisa y pertinente las interacciones en cada tecnología.
- **Inexistencia de un protocolo formal de evaluación y trazabilidad:** Los equipos de estudiantes carecen de un marco de trabajo estandarizado que conecte las decisiones de diseño con pruebas objetivas, lo que eleva el riesgo de construir productos con baja usabilidad, comprometiendo su adopción y permanencia en el mercado.

Esta conjunción de factores fundamenta la necesidad imperativa de establecer un marco metodológico estructurado que asigne y guíe las pruebas de UX según la naturaleza técnica de cada proyecto.

1.1. Objetivos Generales

Diseñar un protocolo de evaluación de experiencia de usuario a través de métricas y escalas de evaluación UX adaptado a múltiples ecosistemas tecnológicos para los diferentes tipos de proyecto que se presentan en las asignaturas que conllevan a la Feria de Software de la UTFSM.

1.2. Objetivos Específicos

- **Investigar** las métricas y escalas estandarizadas de evaluación de Experiencia de Usuario presentes en la literatura técnica, con el propósito de determinar su pertinencia y adaptabilidad frente a diversos entornos tecnológicos.
- **Categorizar** los tipos de proyectos de software basándome en mi propia experiencia en las asignaturas vinculadas a la Feria de Software, estableciendo una taxonomía preliminar basada en los ecosistemas y bases tecnológicas predominantes.
- **Validar y refinar** la taxonomía tecnológica preliminar mediante el levantamiento empírico de información con participantes previos del evento, consolidando las categorías definitivas.
- **Asignar estratégicamente las metodologías y escalas de evaluación de UX** a cada ecosistema tecnológico, estructurando una matriz operativa fundamentada en la triangulación de métricas.
- **Diseñar** un protocolo que articule de manera estructurada las categorías tecnológicas consolidadas con las escalas y métricas de evaluación de UX idóneos para cada entorno.
- **Validación final** mediante la revisión del protocolo con equipos pertenecientes a los proyectos que serán presentados el presente año en la FESW.

Capítulo 2

Estado del arte

En el presente capítulo, con el fin de comprender las distintas aproximaciones existentes para la evaluación de UX, se presenta una revisión del estado del arte. Esta revisión cobra especial relevancia al considerar el contexto de la Feria de Software de la Universidad Técnica Federico Santa María (organizada por el Departamento de Informática), la cual al presente año 2026 acumula más de 30 años de historia. A lo largo de la trayectoria de la FESW, los proyectos presentados han experimentado una notable evolución tecnológica: desde las primeras páginas web estáticas hasta aplicaciones móviles complejas, transitando hoy en día hacia desarrollos impulsados por inteligencia artificial o tecnologías nuevas más complejas.

Frente a este dinamismo, el capítulo reúne investigaciones, metodologías e instrumentos ampliamente utilizados en diferentes contextos tecnológicos, permitiendo identificar los principales criterios, métricas y enfoques empleados para analizar la interacción entre los usuarios y los sistemas. El análisis de estos antecedentes proporciona una base conceptual sólida para fundamentar la selección y aplicación de métodos de evaluación de UX que logren adaptarse eficazmente a las características y bases tecnológicas de cada proyecto.

2.1. Análisis Comparativo de Principios de Usabilidad

Con el propósito de estructurar el marco evaluativo de la presente investigación, resulta indispensable examinar los criterios dictados por los principales referentes teóricos en usabilidad. Para efectos de este trabajo, se contrastarán los postulados de dos pioneros fundamentales en la Interacción Humano-Computadora (HCI): Jakob Nielsen y Ben Shneiderman. El objetivo de este análisis es determinar y seleccionar un único marco de referencia que guíe metodológicamente el protocolo a seguir. Estas directrices operan como estándares de calidad orientados a garantizar que un sistema interactivo sea fácil de aprender, eficiente, seguro y satisfactorio para sus usuarios. En consecuencia, estos marcos no solo guían el diseño de interfaces, sino que proporcionan los criterios fundamentales para auditar y evaluar productos digitales desde la perspectiva de la experiencia de usuario.

Principios determinados por Nilsen

Jakob Nielsen define la usabilidad como un concepto multidimensional que no es una propiedad única, sino que se compone de atributos como la **capacidad de aprendizaje (learnability)**, la **eficiencia**, la **memorabilidad**, una **baja tasa de errores** y la **satisfacción subjetiva** del usuario. Su enfoque se fundamenta en la “Discount Usability Engineering”, que utiliza métodos rápidos y económicos para mejorar las interfaces, destacando la evaluación heurística como una inspección sistemática basada en principios generales. Estas **10 Heurísticas de Usabilidad** [6], listadas en su obra, comienzan con el requisito de un **diálogo simple y natural** (*simple and natural dialogue*) para eliminar información irrelevante y **hablar el lenguaje de los usuarios** (*speak the users’ language*) mediante conceptos familiares en lugar de términos orientados al sistema. Nielsen enfatiza la necesidad de **minimizar la carga de memoria del usuario** (*minimize the users’ memory load*) haciendo que los objetos y opciones sean visibles, y mantener la **consistencia** (*consistency*) para que el usuario no dude sobre el significado de palabras o acciones. El sistema debe proveer siempre **retroalimentación** (*feedback*) sobre su estado y ofrecer **salidas claramente marcadas** (*clearly marked exits*), como una “salida de emergencia” para abandonar estados no deseados sin diálogos extensos.

Además, se deben incluir **atajos** (*shortcuts*) o aceleradores invisibles para expertos, proporcionar **buenos mensajes de error** (*good error messages*) que sean constructivos y claros, y, fundamentalmente, trabajar para **prevenir errores** (*prevent errors*) antes de que ocurran mediante un diseño cuidadoso. Por último, el autor reconoce que puede ser necesaria la **ayuda y documentación** (*help and documentation*) enfocada en las tareas específicas del usuario[6].

Principios determinados por Shneiderman

Ben Shneiderman fundamenta su trabajo en la búsqueda de la **usabilidad universal**, adaptando las interfaces a la diversidad humana en capacidades físicas, cognitivas y culturales para que la tecnología sea accesible para todos. Su filosofía de diseño promueve la **manipulación directa**, permitiendo que la interfaz sea tan fluida que el ordenador “desaparezca” de la conciencia del usuario. Shneiderman establece que el éxito de un sistema debe medirse mediante métricas críticas como el tiempo de aprendizaje, la velocidad de desempeño y la satisfacción subjetiva del operador. Para guiar este proceso, el autor propone sus **8 Reglas de Oro del Diseño de Interfaz** [7], principios refinados durante décadas que comienzan con la instrucción de **esforzarse por la consistencia** (*strive for consistency*) mediante el uso uniforme de terminología, fuentes y colores en todas las pantallas. El diseño debe además **buscar la usabilidad universal** (*seek universal usability*), reconociendo las necesidades de diversos públicos y ofreciendo atajos para expertos, y **ofrecer retroalimentación informativa** (*offer informative feedback*) inmediata ante cada acción. Es vital **diseñar diálogos que generen cierre** (*design dialogs to yield closure*) para dar una sensación de logro al completar tareas, así como centrarse en **prevenir errores** (*prevent errors*) mediante interfaces robustos que ofrezcan instrucciones de recuperación específicas si ocurre un fallo. Complementariamente, el sistema debe **permitir la reversión fácil de acciones** (*permit easy reversal of actions*) para reducir la ansiedad del usuario, **mantener a los usuarios en control** (*keep users in control*) dándoles la seguridad de que ellos manejan el sistema, y, finalmente, **reducir la carga de la memoria a corto plazo** (*reduce short-term memory load*) evitando que deban recordar grandes cantidades de datos entre pantallas[7].

2.1.1. Selección y Fundamentación del Marco Heurístico

A partir de los principios analizados, se determina que las heurísticas de Jakob Nielsen constituyen el marco principal más adecuado para evaluar los proyectos de la Feria de Software de la UTFSM, superando el enfoque de las Reglas de Oro de Ben Shneiderman. Esta decisión se justifica estructuralmente en que Shneiderman concibió sus reglas principalmente como directrices para guiar el proceso de diseño de interfaces. En contraste, Nielsen desarrolló sus heurísticas específicamente como un método de inspección para evaluar la usabilidad de sistemas interactivos ya existentes o en etapas de prototipado funcional, que es precisamente el estado de las soluciones que se presentan en la Feria. Al requerirse un análisis bajo criterios homogéneos y sistemáticos, las heurísticas de Nielsen permiten identificar problemas de usabilidad de manera más rigurosa y estructurada.

Además de su enfoque evaluativo, este marco destaca por su alta versatilidad y validación en la literatura de UX. Su flexibilidad operativa permite examinar objetivamente el abanico diverso de soluciones tecnológicas de los proyectos, que incluye desde aplicaciones móviles y web convencionales hasta plataformas SaaS, herramientas orientadas al emprendimiento y sistemas impulsados por inteligencia artificial. Al ofrecer criterios concretos de inspección, tales como la visibilidad del estado del sistema, la prevención de errores, la consistencia, la eficiencia y el control por parte del usuario, Nielsen facilita una comparación equitativa y estandarizada entre proyectos de distintas bases tecnológicas, asegurando un estándar óptimo para la experiencia del usuario final.

Finalmente, esta rigurosidad cobra especial relevancia en el contexto de la Feria de Software, donde los proyectos no solo buscan la innovación técnica, sino también generar un impacto social y alinearse con los Objetivos de Desarrollo Sostenible (ODS). Para cumplir con este propósito, es fundamental garantizar que las soluciones propuestas sean fáciles de aprender, adoptar y utilizar por sus comunidades objetivo. De este modo, la elección de Nielsen se consolida como la alternativa más apropiada para los fines de este protocolo, proporcionando la metodología de auditoría sistemática que el dinamismo de la Feria de Software requiere.

Si bien las heurísticas de Nielsen proporcionan una base transversal y estandarizada para la inspección de interfaces en etapas tempranas, la evaluación empírica de la experiencia

de usuario requiere un análisis más profundo que considere las limitantes físicas y cognitivas del soporte tecnológico subyacente. La aplicación exclusiva de principios heurísticos resulta insuficiente cuando el proyecto trasciende el software tradicional e incorpora tecnologías emergentes que alteran la naturaleza del control del usuario [6]. Por consiguiente, resulta indispensable examinar cómo la arquitectura del ecosistema tecnológico condiciona la interacción, obligando a adaptar los enfoques de medición empírica antes de definir los instrumentos estandarizados a utilizar.

2.2. Paradigmas de Adaptabilidad y Dependencia Tecnológica en la Evaluación de UX

En esta sección se abordan los fundamentos teóricos y normativos que respaldan la necesidad de diseñar estrategias de evaluación flexibles. A través de la revisión de literatura clásica y contemporánea, se examina cómo la evolución de las plataformas de hardware y software impide la aplicación de métricas universales, exigiendo la adopción de marcos de trabajo maleables y estándares de calidad capaces de ajustarse a las particularidades de cada ecosistema tecnológico.

2.2.1. Determinismo del Soporte Tecnológico y el Contexto de Uso en la Interfaz

A continuación se agrupan algunas de las investigaciones que demuestran cómo la arquitectura física y lógica de un dispositivo (sea móvil, de escritorio, inmersivo o un elemento de hardware puro) condiciona directamente la experiencia del usuario. Aquí se establece que la usabilidad no es una propiedad estática, sino una relación dinámica que obliga a ponderar los atributos de diseño de manera diferenciada según la plataforma elegida.

2.2.1.1. Bases de la ingeniería de usabilidad y dependencia de la plataforma

Se presenta el libro “Usability Engineering” de Jakob Nielsen, quien es ampliamente considerado como el “padre de la usabilidad web” y uno de los pioneros en el mundo de la experiencia de usuario [6]. En su escrito, fundamenta la necesidad de adaptar la evaluación de usabilidad al establecer que esta no es una propiedad unidimensional de la interfaz, sino un atributo relativo que depende de la interacción entre usuarios, tareas y, fundamentalmente, la plataforma tecnológica [6]. En su obra, explica que cada avance, desde los sistemas de comandos hasta las “interfaces de próxima generación” como la Realidad Virtual, altera la naturaleza del control del usuario [6]. Mientras que en una aplicación móvil tradicional el usuario gestiona objetos explícitos a través de una interfaz gráfica, en entornos inmersivos o de “no-comando”, el sistema debe interpretar acciones implícitas, lo que obliga a desplazar el foco de la evaluación desde la eficiencia operativa hacia aspectos como la satisfacción y la presencia [6].

El autor sostiene que los cinco atributos de usabilidad *facilidad de aprendizaje, eficiencia, memorabilidad, gestión de errores y satisfacción* deben ponderarse de manera distinta según la base técnica del proyecto [6]. Por ejemplo, en sistemas de entretenimiento o nuevos medios, la satisfacción subjetiva puede ser el factor dominante, mientras que en sistemas de productividad prima la eficiencia. Nielsen advierte explícitamente que aplicar los mismos métodos y métricas de una tecnología consolidada a una emergente es ineficaz. De este modo, sugiere que las herramientas de ingeniería de usabilidad, como la evaluación heurística, deben sufrir “ligeras modificaciones” para poder capturar los problemas reales y las oportunidades de diseño de cada dominio específico [6].

Finalmente, Nielsen valida que un marco de evaluación monolítico y rígido resulta insuficiente ante la diversidad técnica actual; para él, el éxito de la ingeniería de usabilidad radica en su capacidad de ser un proceso iterativo y adaptable [6]. Es precisamente esta premisa teórica la que justifica y fundamenta la estructuración del protocolo propuesto en este trabajo: un marco estandarizado que no impone una métrica universal, sino que sistematiza la adaptabilidad exigida por Nielsen, dotando a los equipos de desarrollo de una guía formal para seleccionar y aplicar los instrumentos de evaluación UX más pertinentes según su

ecosistema tecnológico [6].

2.2.1.2. La experiencia de usuario en el contexto de los multi-dispositivo

Esta investigación bajo la premisa de Nielsen sobre la relevancia de la plataforma, la investigación contemporánea aborda el desafío del consumo de contenido en múltiples dispositivos, señalando que los usuarios demandan experiencias adaptadas a sus preferencias específicas en cada punto de contacto. El estudio enfatiza que el diseño y la evaluación para entornos multi-dispositivos “es más que simplemente cambiar el tamaño del contenido para mostrarlo en diferentes pantallas”. Esta distinción es crucial, ya que implica que la efectividad de una interfaz depende de proporcionar el contenido correcto en el momento adecuado según la tecnología elegida por el usuario. Para resolver este reto, los autores aplicaron y validaron propuestas integradoras de técnicas y herramientas basadas en la experiencia de usuario (UX) para satisfacer a diferentes grupos en diversos entornos técnicos. El texto concluye que la usabilidad y la accesibilidad deben ser validadas mediante procesos de diseño y evaluación que reconozcan explícitamente las diferencias entre los dispositivos, evitando enfoques universales que no consideran el contexto tecnológico [8].

2.2.1.3. Sinergia entre diseño y capacidades tecnológicas (Software vs. Hardware)

En el artículo “El impacto del design thinking en la tecnología” [9] de la Revista Innovación y Software, examina el impacto del Design Thinking en la tecnología, estableciendo una clara distinción entre el desarrollo de software y el diseño de hardware. Mientras que en el software el enfoque se centra en crear interfaces móviles y plataformas intuitivas, en el diseño de hardware se busca asegurar que los productos físicos sean estéticamente atractivos y funcionalmente aceptables. Este estudio subraya que los productos diseñados bajo este enfoque demuestran una combinación efectiva de estética y usabilidad física, lo que resulta en una experiencia general superior. Un punto clave que conecta con la necesidad de métricas diferenciadas es la afirmación de que la innovación surge de la sinergia entre el enfoque humano y las capacidades tecnológicas específicas de cada producto. De este modo, el texto respalda la premisa de que los marcos de evaluación deben iterar y experimentar basándose

en la naturaleza técnica de la solución, ya sea una aplicación digital o un dispositivo físico [9].

2.2.2. Marcos de Trabajo Flexibles y Estandarización de la Calidad del Software

El propósito de este apartado es fundamentar teóricamente la viabilidad y la necesidad de estructurar un protocolo de evaluación adaptable. Como se estableció en la definición del problema (Capítulo 1), los proyectos desarrollados para la Feria de Software presentan una alta heterogeneidad tecnológica. Ante esta diversidad, resulta imperativo demostrar que el aseguramiento de la calidad del software no se logra mediante la imposición de una métrica única y rígida, sino a través de **marcos de trabajo flexibles** que se ajusten a las restricciones de cada ecosistema. A continuación, se analiza cómo la literatura normativa y de gestión de la Experiencia de Usuario confirma la inexistencia de medidas globales y aboga por estructuras adaptables.

2.2.2.1. Certificación de procesos y estandarización flexible de la calidad

Como antecedente para formalizar estas evaluaciones en la industria bajo estándares de calidad, destaca el trabajo de investigación titulado "Framework para la certificación de procesos de software bajo la norma IRAM-ISO 9001 | ISO-IEC 90003 utilizando metodologías ágiles"[10]. El trabajo propone un marco de trabajo diseñado para guiar a las organizaciones hacia una certificación de procesos de software que sea aplicable independientemente de la metodología ágil puesta en práctica. El objetivo central de esta propuesta es identificar puntos en común entre diversos flujos de trabajo que permitan evaluar la calidad del desarrollo de manera sistemática y rigurosa. La investigación subraya la importancia de diseñar marcos de trabajo que permitan una integración fluida entre las prácticas ágiles de desarrollo y los estándares normativos de evaluación. El estudio concluye que, para lograr una certificación exitosa, es necesario analizar y obtener resultados mediante pruebas que validen la eficiencia real de los procesos implementados. Esta perspectiva es fundamental, ya que apoya la

necesidad de contar con estructuras de evaluación que posean la flexibilidad necesaria para ajustarse a diferentes realidades técnicas, garantizando que el producto final cumpla con los objetivos de calidad y fiabilidad establecidos para el sector tecnológico específico [10].

2.2.2.2. Inexistencia de medidas globales y marcos de trabajo adaptables

En el estudio “Marco de trabajo de evaluación de experiencia de usuario en el desarrollo de software”, de la Revista Cubana de Ciencias Informáticas, se presenta un marco de trabajo diseñado para identificar oportunidades de mejora en el desarrollo de software, partiendo de la premisa de que “no existe una medida global de UX generalmente aceptada”. Las autoras sostienen que la experiencia del usuario es un resultado dinámico que depende del contexto de uso, el cual incluye el rendimiento del sistema y las capacidades funcionales del producto. Uno de los principios rectores de su propuesta es la necesidad de “adaptar el proceso de evaluación de UX a las características del proyecto y el producto”, reconociendo que factores como el tiempo, el presupuesto y el nivel de experiencia técnica determinan la ruta de evaluación más adecuada. La investigación detalla diversas rutas de aplicación que combinan métodos de inspección y de prueba de manera diferenciada según el escenario. Al proponer un “Compendio UX” para categorizar la severidad de las inconsistencias, el texto valida que los criterios de éxito deben ser ajustables y fundamentados en los hallazgos específicos de cada entorno tecnológico [11].

En conclusión, los antecedentes expuestos en esta sección justifican directamente el núcleo de la presente investigación. Al confirmarse la inexistencia de medidas universales para evaluar la experiencia de usuario, se evidencia que intentar auditar todos los ecosistemas tecnológicos de la Feria de Software bajo un único instrumento estandarizado sería un error metodológico. Por consiguiente, la literatura valida que el enfoque correcto para garantizar el control de calidad en las asignaturas formativas es el diseño de un protocolo estructurado capaz de asignar metodologías y métricas específicas según la naturaleza técnica y operativa de cada proyecto.

2.3. Paradigma de Triangulación y Enfoque Plural en la Medición de UX

En esta sección se deja asentado cómo la teoría de Interacción Humano-Computadora (HCI) y los factores humanos respaldan el uso coordinado de métricas post-tarea (fricción inmediata) y post-test (percepción macro) para evaluar sistemas complejos.

2.3.1. El Paradigma de Denzin: Principios de la Triangulación Científica

Históricamente, la evaluación de la usabilidad de un sistema se ha supeditado a la aplicación de un único instrumento global al finalizar la experiencia del usuario. Sin embargo, autores clave en el análisis cuantitativo de UX como Sauro y Lewis[12] postulan que la experiencia es multidimensional y requiere capturar tanto la percepción inmediata de la tarea como la valoración retrospectiva del sistema completo.

Para resolver esta complejidad sin abandonar el enfoque cuantitativo, se recurre a la triangulación intramétodo (within-method triangulation), formalizada originalmente por Denzin (1978)[13]. Este enfoque metodológico defiende que el uso de múltiples escalas psicométricas complementarias sobre un mismo fenómeno permite contrastar constructos cognitivos distintos (como la facilidad operativa versus el confort físico o la confianza), mitigando los sesgos de memoria y recencia, y otorgando una validez convergente superior al diseño experimental. Este enfoque postula que la evaluación de sistemas interactivos complejos no debe depender de un único instrumento declarativo, sino de la combinación estratégica de herramientas complementarias dentro de un mismo diseño operativo. Al coordinar métricas de diferente espectro temporal (micro-tácticas y macro-estratégicas) o de distinta naturaleza funcional (usabilidad de software frente a ergonomía de hardware), la triangulación permite mitigar sesgos críticos como el enmascaramiento lúdico en entornos educativos o la tolerancia psicofisiológica en interfaces inmersivas, garantizando un estándar de auditoría riguroso y replicable para plataformas de bases tecnológicas heterogéneas [13].

En consecuencia, para que la triangulación metodológica sea efectiva ante la diversidad de ecosistemas tecnológicos, el diseño de evaluación empírica debe considerar simultáneamente tres aristas fundamentales de la interacción:

- **Arista Temporal (Micro-táctica vs. Macro-estratégica):** Considera la necesidad de capturar la experiencia en dos momentos psicológicos distintos. Se debe evaluar tanto la fricción cognitiva inmediata generada durante la ejecución de una tarea específica (métrica post-tarea) como la percepción retrospectiva y global del sistema al finalizar la sesión (métrica post-test) [12].
- **Arista Funcional (Usabilidad de Software vs. Factores de Frontera):** Reconoce que la experiencia trasciende la interfaz gráfica. Se debe contrastar el rendimiento del entorno digital con las limitantes físicas o psicológicas impuestas por el hardware y el contexto de uso, tales como el confort corporal en dispositivos vestibles, la cinetosis en simuladores inmersivos o la confianza algorítmica en sistemas automatizados [6].
- **Arista de Fiabilidad (Percepción Subjetiva vs. Auditoría Objetiva):** Contempla el cruce entre la declaración de satisfacción del usuario final y las métricas de rendimiento analítico o evaluación experta. Esta arista es crítica para aislar problemas técnicos del entorno (ruido, latencia de sensores) de las deficiencias propias del diseño interactivo.

Al estructurar la evaluación en torno a estas tres aristas, se mitigan sesgos como el enmascaramiento lúdico o la tolerancia psicofisiológica, garantizando que las particularidades tecnológicas de cada proyecto sean auditadas de manera integral y preparando el terreno para una asignación modular de métricas.

2.4. Escalas de evaluación para la medición de la UX según su entorno tecnológico

Hoy en día, la experiencia de usuario ha adquirido una relevancia crítica; sin embargo, la coexistencia de proyectos basados en naturalezas tecnológicas tan diversas hace que adoptar un método de evaluación único e universal sea una tarea inverosímil. Ante este escenario, y en concordancia con el paradigma de triangulación metodológica expuesto previamente, resulta fundamental identificar un catálogo de métricas especializadas que puedan ser combinadas estratégicamente según las restricciones y capacidades de cada ecosistema tecnológico.

A pesar de esta necesidad, en la literatura especializada de UX no existe un consenso ni una taxonomía estandarizada que asigne de manera taxativa qué evaluación corresponde a cada entorno. Persiste un vacío operativo en la industria al momento de auditar soluciones complejas. Por ello, este apartado presenta un levantamiento y análisis de diversas escalas e instrumentos psicométricos validados, clasificados según su foco de aplicación dominante (dispositivos vestibles, videojuegos, simuladores inmersivos, sistemas críticos, entre otros). A través de esta revisión, se examinan herramientas modulares diseñadas para cuantificar dimensiones específicas como la usabilidad transaccional, la carga cognitiva, el confort físico y la sensación de presencia; sentando las bases teóricas para la futura articulación de un protocolo de evaluación multidimensional.

Para efectos de la presente revisión bibliográfica, y con el propósito de exponer los instrumentos descubiertos de manera estructurada, se ha adoptado una categorización de entornos tecnológicos derivada directamente de los dominios de aplicación más recurrentes en la literatura contemporánea sobre Interacción Humano-Computadora y factores humanos. Es fundamental aclarar que esta división metodológica que abarca desde dispositivos vestibles hasta entornos inmersivos no constituye la taxonomía definitiva del presente proyecto, sino que refleja los principales nichos de investigación donde la academia ha focalizado el desarrollo empírico de métricas de usabilidad. Esta organización referencial permite presentar cada escala psicométrica estandarizada dentro de su contexto de origen ¹, estableciendo así el catálogo de herramientas necesario para la posterior formulación empírica de los ecosistemas locales abordados en esta investigación.

2.4.1. Aplicaciones Web

A continuación se presenta un texto sugerido para la escala de evaluación para medir la UX en páginas web/desktop:

2.4.1.1. Escala de Usabilidad del Sistema

John Brooke desarrolló System Usability Scale (SUS)[14] como una herramienta de “bajo costo” y “rápida y sucia” (quick and dirty) para realizar evaluaciones globales de la usabilidad de sistemas en contextos industriales. El autor sostiene que la usabilidad no existe de forma absoluta, sino que es una cualidad de la “adecuación a un propósito” definida por el contexto específico: los usuarios, las tareas y el entorno. La escala SUS consiste en 10 ítems tipo Likert (desde "Totalmente en desacuerdo" hasta "Totalmente de acuerdo") que permiten recolectar evaluaciones subjetivas de los usuarios tras haber interactuado con un sistema. Para evitar sesgos de respuesta, Brooke diseñó el cuestionario alternando sentencias positivas y negativas, obligando al usuario a leer cada afirmación antes de marcar su grado de acuerdo o desacuerdo. El resultado final es un puntaje compuesto que va de 0 a 100, el cual proporciona una visión general de la usabilidad percibida y es lo suficientemente robusto para comparar diferentes versiones de un producto o sistemas competidores [14].

2.4.2. Aplicaciones móviles

A continuación se presenta un texto sugerido para la escala de evaluación para medir la UX en Aplicaciones móviles:

2.4.2.1. Standardized User Experience Percentile Rank for Mobile

Como respuesta a las limitaciones de aplicar cuestionarios genéricos de sitios web en aplicaciones portátiles, Sauro desarrolló la escala **SUPR-Qm** (*Standardized User Experience Percentile Rank for Mobile*) [16]. Este instrumento psicométrico estandarizado consta de

13 reactivos diseñados explícitamente para evaluar la experiencia de usuario en aplicaciones móviles nativas e híbridas. La escala se fundamenta en un modelo multidimensional que evalúa cuatro factores clave de la interacción en dispositivos inteligentes: **usabilidad percibida** (facilidad y eficiencia), **apariencia estética** (calidad del diseño gráfico en pantalla reducida), **confianza/seguridad** (credibilidad en el manejo de transacciones y datos sensibles) e **intención de lealtad** (medida mediante el índice *Net Promoter Score* o NPS). La incorporación de esta herramienta en un marco de evaluación permite obtener un diagnóstico macro de la calidad del software móvil que trasciende la usabilidad funcional pura, correlacionando el rendimiento técnico con la percepción de marca y la seguridad del operador. El cuestionario se aplica solicitando al usuario que evalúe las 12 primeras afirmaciones (asociadas a usabilidad, confianza y atractivo visual) mediante una escala Likert de 5 puntos, desde 1 ('Totalmente en desacuerdo') hasta 5 ('Totalmente de acuerdo'). Finalmente, el ítem número 13 (lealtad/NPS) se califica pidiendo al usuario que indique la probabilidad de recomendar la plataforma en una escala numérica del 0 al 10 [16].

2.4.3. Videojuegos

A continuación se presenta un texto sugerido para la escala de evaluación para medir la UX en Videojuegos:

2.4.3.1. The Game User Experience Satisfaction Scale (GUESS): A Psychometric Validation

La escala **GUESS-18** [17] es una versión abreviada y validada de la *Game User Experience Satisfaction Scale (GUESS)*, diseñada específicamente para evaluar de manera rápida y exhaustiva la satisfacción del usuario en videojuegos en general, sin limitarse a un género o plataforma de juego en particular. Fue desarrollada mediante un proceso psicométrico riguroso que incluyó análisis factoriales exploratorios y confirmatorios para asegurar que, a pesar de su brevedad, mantuviera la capacidad de medir las dimensiones críticas de la experiencia de juego identificadas en la versión original de 55 ítems. Esta versión de 18 ítems

mantiene las nueve subescalas fundamentales del instrumento original: **Usabilidad, Narrativa, Absorción, Disfrute, Libertad Creativa, Estética de Audio, Gratificación Personal, Conectividad Social y Estética Visual**. Su principal ventaja es la aplicabilidad práctica en entornos industriales de *playtesting*, donde permite obtener retroalimentación valiosa de diversos géneros de juegos y perfiles de jugadores en un tiempo significativamente menor, facilitando la toma de decisiones en el diseño y desarrollo de videojuegos. La evaluación se lleva a cabo mediante un cuestionario donde el participante responde a 18 afirmaciones utilizando una escala Likert de 7 puntos, con anclajes que van desde 1 ('Totalmente en desacuerdo') hasta 7 ('Totalmente de acuerdo')[17].

2.4.4. Realidad Virtual y Simulación

A continuación se presenta un texto sugerido para la escala de evaluación para medir la UX en AR/VR:

2.4.4.1. Measuring Presence in Virtual Environments: A Presence Questionnaire

El Cuestionario de Presencia (**PQ**) [18] fue diseñado para medir el grado en que los individuos experimentan la sensación de “estar allí” dentro de un entorno virtual (VE). Esta herramienta se basa en factores teóricos como el control, el realismo y la distracción, permitiendo cuantificar la intensidad de la experiencia subjetiva a través de un formato de autoinforme. Mediante análisis de conglomerados, se identificaron tres subescalas principales que estructuran la evaluación: **Involucramiento/Control**, que mide la capacidad de actuar y la respuesta del entorno; **Naturalidad**, que evalúa qué tan coherente es la experiencia con la realidad; y **Calidad de la Interfaz**, que observa si los dispositivos de control distraen al usuario de su tarea. Los resultados han demostrado que existe una relación positiva entre la presencia y el rendimiento en la tarea, así como una correlación negativa con el malestar por simulación (SSQ). Su aplicación consiste en un cuestionario de 32 ítems que el usuario debe responder de forma independiente tras finalizar una tarea en el entorno virtual. Cada reactivo se evalúa utilizando una escala Likert de 7 puntos con anclajes semánticos que varían según la pregunta (por ejemplo, desde 1 “Para nada/En absoluto” hasta 7 “Completamente”),

permitiendo medir detalladamente el grado de inmersión [18].

2.4.4.2. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness

El Cuestionario de Malestar por Simulación (SSQ) [19] se presenta como un método mejorado para cuantificar los síntomas de malestar que surgen específicamente en simuladores visuales de alta fidelidad. A diferencia de las escalas diseñadas para el mareo común por movimiento (MS), el SSQ se enfoca en el malestar por simulación (SS), el cual suele ser menos severo y originarse por distorsiones visuales o asincronías en la imagen. Esta herramienta se derivó de un cuestionario previo tras eliminar síntomas irrelevantes y aplicar análisis factoriales para lograr una medición más precisa y adaptable a sistemas computarizados. La escala organiza los síntomas en tres grupos o subescalas: **Oculomotores** (como fatiga ocular y dolor de cabeza), **Desorientación** (que incluye mareos y vértigo) y **Náuseas** (que abarca conciencia estomacal y salivación aumentada). Esta estructura permite un diagnóstico detallado sobre qué aspecto del simulador está causando problemas, permitiendo distinguir si las fallas provienen del sistema visual o de deficiencias más generalizadas. El proceso de calificación genera puntajes para cada subescala y un índice de severidad total, lo que facilita el control de calidad de los dispositivos. El instrumento se administra solicitando al usuario que indique la presencia e intensidad de 16 síntomas fisiológicos específicos en el momento posterior a la experiencia inmersiva. Para ello, el participante utiliza una escala de 4 puntos: 0 (Ausente), 1 (Leve), 2 (Moderado) y 3 (Grave). En la práctica, el SSQ se utiliza para monitorear el desempeño de los simuladores y detectar errores de calibración o problemas mecánicos mediante el reporte de los usuarios. También es una herramienta clave para estudiar cómo los sujetos se adaptan al estrés del simulador tras exposiciones repetidas, demostrando que el malestar suele disminuir significativamente después de varias sesiones. Al establecer puntuaciones base, las organizaciones pueden evaluar el impacto de modificaciones de ingeniería y tomar precauciones para reducir riesgos en actividades posteriores [19].

2.4.5. Dispositivos Vestibles (Wearables)

A continuación se presenta un texto sugerido para la escala de evaluación para medir la UX en dispositivos vestibles:

2.4.5.1. Assessing the Comfort of Wearable Computers

La Escala de Calificación de Confort (**CRS**)[20] es una herramienta multidimensional diseñada para evaluar la usabilidad de las computadoras vestibles (wearables), basándose en la premisa de que un único valor global es insuficiente para medir el confort de manera apropiada. El sistema analiza seis dimensiones fundamentales: **emoción, sujeción (attachment), daño, percepción de cambio, movimiento y ansiedad**. Estas categorías permiten capturar no solo el malestar físico o la restricción de movimiento, sino también el impacto psicológico del uso del dispositivo. Esta escala fue desarrollada mediante un escalamiento multidimensional de términos asociados al confort. Su aplicación práctica sirve como una guía para diseñadores e investigadores, permitiéndoles resaltar diferencias específicas entre dispositivos y determinar la efectividad de las modificaciones realizadas en el diseño ergonómico para mejorar la experiencia del usuario. Para su ejecución, se requiere que el participante complete una rutina activa con el dispositivo (estado dinámico), nunca en estado sedentario. Inmediatamente después, el usuario califica su percepción en cada una de las seis dimensiones mediante una escala continua que va desde 0 (confort perfecto) hasta 20 (disconfort máximo) [20].

2.4.6. Sistemas Automatizados e Inteligencia Artificial

A continuación se preseta un texto sugerido para proyectos que implementan automatización e IA:

2.4.6.1. Foundations for an Empirically Determined Scale of Trust in Automated Systems

La Escala **Jian Trust** [22] aborda la necesidad crítica de medir de forma efectiva la confianza de los usuarios en los sistemas automatizados para comprender la relación entre la percepción del humano y el uso de sistemas computarizados. El trabajo critica que las herramientas de evaluación previas se basaban en nociones teóricas generales y no en análisis empíricos que permitieran distinguir entre la confianza humano-humano y la confianza humano-máquina. Los autores identificaron 12 factores específicos que componen **la confianza en la automatización, incluyendo la fiabilidad, la seguridad, la integridad y la familiaridad**. Un hallazgo fundamental es que el desarrollo de una escala diferenciada proporciona una herramienta más válida y fiable para evaluar sistemas automatizados específicos, reconociendo que la usabilidad y la aceptación dependen de cómo la tecnología es percibida por el operador. La investigación concluye que la confianza y la desconfianza operan como opuestos en una misma dimensión, lo que permite simplificar el diseño de los instrumentos de medición. La prueba se administra a través de un cuestionario compuesto por 12 afirmaciones (alternando sentencias positivas y negativas), donde el operador califica la intensidad de su percepción hacia el sistema automatizado utilizando una escala Likert de 7 puntos, estableciendo el anclaje 1 como “Nada en absoluto” y 7 como “Extremadamente”. Además, el estudio enfatiza que los factores de confianza pueden ser dinámicos y cambiar a medida que el sistema se degrada o mejora, lo que refuerza la importancia de contar con marcos de evaluación que se ajusten a la realidad tecnológica del producto. Este enfoque respalda la necesidad de utilizar métricas especializadas que capturen las particularidades de la interacción hombre-máquina, validando que el contexto tecnológico define los criterios de éxito de la experiencia de usuario [22].

2.4.7. Métricas Transversales y Marcos Complementarios de UX

A continuación se presentan texto sugerido sobre las escalas globales, herramientas post-tarea, métodos de auditoría analítica por expertos y fundamentos estadísticos que se pueden adaptar a la naturaleza tecnológica de la interfaz.

2.4.7.1. Comparison of Three One-Question, Post-Task Usability Questionnaires

La escala **SEQ** (Single Ease Question) [23], funciona como una métrica de evaluación de usabilidad de tipo pos-tarea, diseñada para medir la percepción del usuario inmediatamente después de finalizar una actividad específica. Teóricamente, este instrumento se fundamenta en una única pregunta que solicita al participante calificar qué tan difícil o fácil le resultó completar la tarea, lo que permite obtener información diagnóstica con una alta validez al capturar la experiencia del usuario en el momento exacto en que ocurre. Su implementación más común y confiable exige una regla operativa temporal estricta: se debe presentar la pregunta al usuario en el instante exacto en que termina o abandona un escenario de tarea, solicitándole que califique la dificultad de la acción mediante una escala Likert lineal de 7 puntos (donde 1 representa “Muy difícil” y 7 “Muy fácil”). , permitiendo a los investigadores registrar los sentimientos del usuario de manera precisa sin imponer una carga cognitiva excesiva. A pesar de su extrema brevedad, la SEQ ha demostrado ser una herramienta robusta y sensible que se correlaciona significativamente con métricas de rendimiento como el tiempo de la tarea, las tasas de finalización y los errores, siendo especialmente valorada en la industria por su facilidad de administración y su capacidad para predecir la usabilidad general [23].

2.4.7.2. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research

El **NASA-TLX** (NASA-Task Load Index) [24] es una técnica de calificación multidimensional desarrollada para proporcionar una estimación sensible y confiable de la carga de trabajo subjetiva. Esta herramienta define la carga como el “costo” en el que incurre un operador humano para alcanzar un nivel específico de desempeño, surgiendo de la interacción entre los requisitos de la tarea y las percepciones, habilidades y comportamientos del individuo. Para realizar esta medición, la escala evalúa seis factores clave: demanda mental, demanda física, demanda temporal, rendimiento, esfuerzo y nivel de frustración. A diferencia de las escalas de puntuación única, el NASA-TLX utiliza un promedio ponderado basado en la importancia

relativa que el usuario asigna a cada factor para una tarea específica. Debido a su alta capacidad diagnóstica y facilidad de aplicación, se considera un instrumento práctico y superior para evaluar el impacto de diversas actividades en entornos operativos bajo alta presión. El procedimiento de evaluación se ejecuta en dos fases, inmediatamente después de finalizada una tarea crítica. En la primera fase **ponderación**, el usuario realiza 15 comparaciones por pares entre las dimensiones para definir cuál representó una mayor fuente de carga. En la segunda fase **calificación**, el participante marca el nivel de intensidad experimentado en cada dimensión sobre una escala continua que va de 0 a 100 puntos, subdividida en incrementos de 5 unidades, con anclajes semánticos en los extremos [24].

2.4.7.3. Evaluación PURE (Practical Usability Rating by Experts)

La evaluación **PURE** [25] es una técnica analítica (en lugar de empírica) diseñada para cuantificar con alta precisión la fricción cognitiva y el esfuerzo físico que un usuario típico encontraría al usar una interfaz. A diferencia de las pruebas con usuarios, PURE se basa en el juicio de expertos familiarizados con principios de UX y heurísticas universales (Usability Universal), lo que la hace ideal para proyectos con limitaciones de tiempo o dificultad para reclutar participantes especializados. PURE se considera una extensión lógica del recorrido cognitivo (cognitive walkthrough) y la evaluación heurística, añadiendo una métrica cuantitativa para predecir la facilidad de la tarea. La evaluación se ejecuta de forma analítica, requiriendo que el equipo desarrollador descomponga previamente el flujo transaccional crítico en pasos secuenciales discretos. Posteriormente, el experto auditor evalúa la interfaz sin involucrar usuarios finales, asignando a cada paso una puntuación de fricción de 1 a 3 (1 = baja fricción/carga nula, 2 = fricción media, 3 = fricción alta/bloqueo crítico) [25].

2.4.7.4. UMUX-LITE – When There’s No Time for the SUS

La escala **UMUX-LITE** [26] funciona como una herramienta de evaluación de usabilidad extremadamente simplificada, compuesta por solo dos ítems de tono positivo que miden las dimensiones fundamentales de utilidad (“Las capacidades de este sistema cumplen con

mis requisitos”) y facilidad de uso (“Este sistema es fácil de usar”). Teóricamente, este instrumento se fundamenta en el Modelo de Aceptación Tecnológica y se presenta como una alternativa ágil a la escala SUS de 10 ítems para situaciones de tiempo limitado o encuestas de gran tamaño. Utiliza un ajuste estadístico de regresión para asegurar que sus resultados finales sean equivalentes y comparables con los estándares globales de la métrica SUS. La escala se aplica de manera posterior a la prueba (post-test) presentando al usuario las dos únicas afirmaciones. El participante debe reflejar su percepción subjetiva marcando su grado de acuerdo en una escala Likert de 7 puntos, anclada en 1 (“Totalmente en desacuerdo”) y 7 (“Totalmente de acuerdo”) [26].

2.4.7.5. Quantifying the User Experience: Practical Statistics for User Research

El manual *Quantifying the User Experience: Practical Statistics for User Research* [12] funciona como un marco metodológico transversal. Por consiguiente, **no está restringido a un tipo de proyecto o ecosistema tecnológico en particular, sino que está considerado para ser aplicado en cualquier desarrollo interactivo (aplicaciones web, móviles, hardware o entornos inmersivos) que requiera validación empírica y justificación estadística de sus métricas de usabilidad.** Es una guía exhaustiva diseñada para que investigadores y profesionales puedan cuantificar, calificar y justificar sus datos de experiencia de usuario con confianza. El texto opera como un manual práctico para resolver problemas cuantitativos comunes, respondiendo a interrogantes críticas como si un producto es más usable que la competencia o si se puede asegurar estadísticamente que un porcentaje específico de usuarios completará una tarea en su primer intento. La obra proporciona una base sólida en teorías estadísticas, extrayendo conocimientos de décadas de literatura sobre factores humanos, ingeniería industrial y psicología. En su segunda edición, el libro incluye un capítulo dedicado al modelado lineal general, que abarca la correlación, la regresión y el análisis de varianza (ANOVA). Esta última técnica estadística cobra vital importancia metodológica en proyectos de alta complejidad física, como los dispositivos vestibles (*wearables* 2.4.5), donde es mandatorio aplicar un ANOVA de medidas repetidas para evaluar las fluctuaciones del confort ergonómico a lo largo del tiempo. Para facilitar la ejecución de estos análisis, los autores

ofrecen soluciones concretas mediante fórmulas de hojas de cálculo y calculadoras especializadas, permitiendo a los equipos de desarrollo procesar métricas de agregación dinámicas y comunicar sus hallazgos de manera objetiva, independientemente de la base tecnológica evaluada [12].

2.5. Consolidación de instrumentos de medición según entorno tecnológico

A modo de síntesis, la revisión del estado del arte evidencia la existencia de un amplio y diverso catálogo de instrumentos psicométricos y analíticos para la evaluación de la Experiencia de Usuario. Tal como se ha fundamentado a lo largo de este capítulo, la adopción de una métrica universal resulta inviable frente a la heterogeneidad tecnológica de las plataformas actuales. La base tecnológica de cada proyecto condiciona intrínsecamente los factores críticos de interacción, requiriendo instrumentos especializados que capturen desde la usabilidad transaccional hasta el desgaste fisiológico, la carga cognitiva o la confianza algorítmica. Por consiguiente, con el propósito de consolidar y delimitar las herramientas que efectivamente conformarán el modelo de evaluación propuesto en esta investigación, la Tabla 2.1 presenta un resumen comparativo de las escalas seleccionadas. Estas herramientas han sido ordenadas de acuerdo con los entornos tecnológicos definidos previamente en la literatura, detallando el objetivo principal y las dimensiones específicas de medición de cada instrumento.

Tabla 2.1: Resumen comparativo de escalas de evaluación UX según categoría tecnológica para la cual fueron creadas.

Categoría Tecnológica	Escala	Objetivo Principal	¿Qué mide?
Aplicaciones Web	SUS	Evaluar usabilidad retrospectiva de manera rápida y global.	Percepción subjetiva de facilidad de aprendizaje y facilidad de uso del sistema.
Aplicaciones Móviles	SUPR-Qm	Realizar un diagnóstico holístico de la calidad de interfaces en dispositivos inteligentes.	Usabilidad, apariencia estética, confianza/seguridad e intención de lealtad (NPS).
Videojuegos	GUESS-18	Medir de forma rápida y exhaustiva la satisfacción del jugador frente al sistema.	Usabilidad, narrativa, ensimismamiento, disfrute, libertad creativa, estéticas, gratificación y conectividad.
Realidad Virtual y Simulación	PQ	Cuantificar la inmersión psicológica del usuario en el entorno inmersivo.	Involucramiento/control, naturalidad de interacción y calidad de la interfaz (distracciones).
	SSQ	Diagnosticar el desgaste fisiológico derivado del hardware o el renderizado.	Cinetosis a través de factores oculomotores, desorientación espacial y náuseas.
Dispositivos Vestibles (Wearables)	CRS	Evaluar de manera multidimensional el confort físico y el impacto psicológico.	Emoción, sujeción, daño físico, cambio percibido, movimiento y ansiedad.

Sistemas Automatizados e Inteligencia Artificial	Jian Trust Scale	Medir empíricamente la confianza del usuario en la toma de decisiones del algoritmo.	Constructo unidimensional compuesto por la confianza activa frente a la desconfianza/engaño del sistema.
Métricas Transversales y Complementarias	SEQ	Medir la fricción cognitiva y táctica inmediata generada en una tarea específica.	Percepción de facilidad o dificultad justo al momento de completar o abandonar un flujo.
	NASA-TLX	Proporcionar una estimación sensible de la carga de trabajo subjetiva bajo presión.	Exigencias mentales, físicas, temporales, nivel de rendimiento, esfuerzo invertido y frustración.
	PURE	Predecir métricas de usabilidad mediante inspección heurística de descuento analítica.	Fricción cognitiva y esfuerzo físico impuesto por la interfaz, auditado paso a paso por expertos.
	UMUX-LITE	Ejecutar una evaluación ágil de usabilidad en escenarios con restricción de tiempo.	Limitado exclusivamente a la percepción de utilidad de las funciones y facilidad de uso global.

Capítulo 3

Propuesta de Solución

A partir de la problemática identificada en las asignaturas que concluyen en la Feria de Software de la UTFSM, donde se observa una postergación sistemática de los atributos de usabilidad frente a la urgencia de la implementación técnica, se diseñará un protocolo de evaluación adaptado a múltiples ecosistemas tecnológicos. Como se expuso en el estado del arte, carece en la literatura de un marco metodológico estandarizado que categorice las soluciones de software según su naturaleza operativa y, simultáneamente, asigne métricas o evaluaciones específicas para cada entorno. Por consiguiente, para estructurar una solución frente a esta brecha, resulta imperativo establecer inicialmente una taxonomía que clasifique los desarrollos de la feria en función de sus bases tecnológicas predominantes y sus paradigmas de interacción de cara al usuario final.

Para llevar a cabo la propuesta de solución, la metodología se ha estructurado en las siguientes fases secuenciales:

- **Paso 0: Investigación inicial.** Se realizó una revisión de las distintas metodologías y métricas de evaluación de Experiencia de Usuario existentes en la literatura técnica (Capítulo 2).
- **Paso 1: Conceptualización y categorización preliminar.** Sistematizar los proyectos de los 30 equipos que cursaron las asignaturas en 2024 (12 de Campus Casa Central y 18 de Campus San Joaquín) e integrar el conocimiento empírico adquirido por la

autora como estudiante y ayudante, estableciendo una taxonomía detallada basada en los ecosistemas tecnológicos y entornos de desarrollo específicos en los que fueron ejecutados.

- **Paso 2: Validación de la categorización preliminar.** Se realizó una encuesta abierta dirigida a estudiantes que presentaron sus proyectos en las Ferias de Software realizadas entre 2019 y 2025.
- **Paso 3: Reestructuración de las categorías.** Se realizó una reformulación en base a los resultados obtenidos en la encuesta de validación de categorías, con el fin de refinar y establecer las categorías definitivas que serán utilizadas.
- **Paso 4: Asignación estratégica de metodologías y métricas de UX.** Creación de una matriz operativa que articula cada ecosistema con una dupla validada de escalas psicométricas (triangulación metodológica).
- **Paso 5: Diseño del protocolo de evaluación UX para proyectos de la FESW.** Se creó el protocolo integrando las categorías creadas y las evaluaciones de UX investigadas, definición de los lineamientos de ejecución y creación del marco de evaluación (UX-Scorecard) para su aplicación práctica dentro del ciclo de las asignaturas.
- **Paso 6: Validación del protocolo.** se realizó un análisis del protocolo creado mediante una encuesta y la aplicación de testing a los estudiantes de INF-228 de 2026

3.1. Paso 1: Conceptualización y categorización preliminar

Como primera fase de la metodología, se llevó a cabo un proceso de conceptualización con el objetivo de establecer una clasificación preliminar de los ecosistemas de software. Para ello, se analizó una muestra representativa de 30 proyectos presentados en la Feria de Software de la edición 2024, de los cuales 12 pertenecieron al Campus Casa Central y 18 al Campus San Joaquín. La selección de esta cohorte específica y la estructuración de la clasificación obedecen a un muestreo de carácter intencional, fundamentado en la participación directa y experiencial de la autora, tanto en calidad de estudiante de las asignaturas involucradas como

en su posterior rol de ayudante académica. Este doble criterio permite que el análisis empírico se construya desde una perspectiva de observación participante y longitudinal, dotando a la investigación de un conocimiento profundo y de primera fuente sobre el contexto de desarrollo, las dinámicas de trabajo, la evolución de los requerimientos a lo largo del tiempo y las exigencias tecnológicas reales enfrentadas por los equipos.

Posteriormente, esta revisión sistemática tuvo como propósito identificar las bases tecnológicas predominantes en las propuestas de los estudiantes. La categorización se estructuró tomando como criterio central la tecnología principal utilizada de cara a la interacción con el usuario final, evaluando para este fin las descripciones funcionales y técnicas declaradas en los portales web oficiales de cada proyecto.

3.1.1. Análisis de proyectos presentados en la Feria de Software 2024 y tecnología predominante de cada proyecto:

Tras el análisis realizado a los 30 proyectos del 2024 se lograron obtener 6 categorías distintas a las cuales se decidió agregar 2 más bajo un análisis profundo e integrar el conocimiento empírico adquirido por la autora como estudiante y ayudante sobre proyectos posteriores de la FESW y la realidad tecnológica actual, con esto se logró un total de 8 categorías preliminares donde se incluye la descripción de cada una y se enlistan los proyectos pertenecientes a ellas con una breve descripción de cada uno.

Para estructurar esta clasificación y comprender la naturaleza de cada proyecto, es imperativo establecer que, en el contexto de esta investigación, el término “ecosistema” no se restringe exclusivamente al soporte físico o hardware utilizado (como ocurre de manera evidente en la Realidad Virtual o los dispositivos Wearables). Por el contrario, se define como un entorno integral de desarrollo que condiciona de manera única la interacción del usuario final.

Bajo esta premisa, la taxonomía propuesta adopta un enfoque multidimensional, clasificando las soluciones de software en base a dos dimensiones fundamentales:

- **Dimensión de Soporte Físico:** Agrupa los proyectos donde la experiencia de usuario está dictada por las capacidades y restricciones del dispositivo de hardware (ej. interfaces táctiles en móviles, ergonomía en wearables o cinetosis en visores inmersivos).

- **Dimensión de Arquitectura Lógica y Contexto Operativo:** Agrupa los proyectos donde, independientemente del hardware tradicional utilizado (como un monitor de escritorio), la complejidad de los datos, la inmediatez requerida y la carga cognitiva definen un entorno crítico propio (ej. toma de decisiones bajo estrés en Ciberseguridad o interpretación de modelos de Inteligencia Artificial en plataformas de Big Data).

Esta distinción permite que proyectos que carecen de un hardware físico dedicado, pero que poseen exigencias operativas críticas, sean categorizados de manera independiente. De este modo, se justifica la necesidad de asignar metodologías y métricas de evaluación UX especializadas que logren capturar la fricción real en cada uno de estos escenarios.

Nota Conceptual: Alcance del término “Transaccional”

En el contexto de la presente taxonomía y de la Interacción Humano-Computadora, el concepto de sistema, flujo o servicio **transaccional** no se restringe a operaciones financieras, monetarias o pasarelas de pago. En ingeniería de software, una *transacción* se define como cualquier intercambio estructurado de información entre el usuario y el sistema que altera el estado de la base de datos o el modelo de negocio [4]. Por consiguiente, registrar un paciente en un sistema de salud, completar un formulario de varias etapas, subir un inventario o emitir una alerta de seguridad, son considerados flujos transaccionales complejos.

1. Ecosistema de Software de Gestión (Web / Desktop)

Software orientado a la productividad, tareas administrativas o transaccionales en pantallas grandes, incluyendo plataformas SaaS y análisis de datos.

- **AuditForge:** Herramienta de auditoría de código abierto que automatiza la generación de informes de seguridad y detección de vulnerabilidades.
- **EmergTrack:** Sistema enfocado en la automatización y análisis para el registro de insumos médicos en centros de urgencias.

- **VentriFLOW:** Software que utiliza IA para la segmentación automática de imágenes, permitiendo visualización 3D y simulación de flujo sanguíneo cardíaco.
- **Molecular Craft:** Plataforma especializada para la investigación farmacéutica (CADD) con herramientas de docking y visualización molecular.
- **GeoTMM:** Algoritmo de Map Matching impulsado por IA para optimizar el análisis geoespacial y la precisión de rutas viales.
- **MineGuard:** Sistema para la gestión y seguimiento en tiempo real de camiones mineros, proporcionando datos estadísticos de operación.
- **QAGuardIAN (PanDev):** Solución de IA que automatiza la revisión de ventas telefónicas para detectar errores y cumplir normativas.
- **IATU (IA Testing User):** IA diseñada para analizar videos de pruebas de usuarios, optimizando la usabilidad y accesibilidad de sitios web.
- **EduvidIA:** Herramienta de creación acelerada de contenido educativo mediante videos de animación whiteboard generados por IA.
- **AI Want 2 Teach:** Plataforma que utiliza IA para optimizar la revisión de preguntas y apoyar el proceso de enseñanza-aprendizaje superior.
- **Digital Skill App:** Aplicación web para evaluar competencias digitales ciudadanas mediante simulaciones prácticas basadas en DigComp.
- **MarIAAna:** SaaS que emplea cálculos matemáticos e IA para reducir el área de búsqueda en operaciones de rescate oceánico (Información externa proporcionada por el usuario).

2. Ecosistema de Servicios Móviles (Mobile Tradicional)

Aplicaciones móviles diseñadas para la portabilidad y uso en movimiento.

- **Refracted Speech:** App que ofrece prediagnóstico y ejercicios personalizados para personas con disartria.

- **ÚneteApp:** Plataforma colaborativa gratuita para el descubrimiento y promoción de actividades comunitarias y prescripción social.
- **SIGO:** Aplicación de seguridad personal y seguimiento en tiempo real.
- **Ranner:** Aplicación móvil para conectar usuarios y explorar nuevas formas de aprendizaje interactivo.
- **sandiApp:** Aplicación de seguimiento nutricional que conecta pacientes con nutricionistas mediante un asistente virtual (Información externa proporcionada por el usuario).
- **SentIAlert:** Aplicación móvil de seguridad con IA para detectar situaciones de riesgo y alertar protectores (Información externa proporcionada por el usuario).

3. Ecosistema de Captura y Visión (Mobile de Captura)

Software donde la cámara y el procesamiento de imágenes son el núcleo de la funcionalidad.

- **DialogApp:** Aplicación que traduce en tiempo real y de forma bidireccional entre Lengua de Señas Chilena (LSCh) y español.
- **MedIntification:** Solución que utiliza captura de fotos para acelerar y automatizar el inventario de insumos médicos.
- **WakeUp Driver:** Aplicación que monitorea al conductor mediante la cámara frontal para detectar fatiga y emitir alertas.
- **Geoviality:** Plataforma que utiliza procesamiento de imágenes e IA para detectar y clasificar irregularidades en la calzada.

4. Ecosistema de Entretenimiento (Videojuegos)

Proyectos orientados al desafío, gamificación o educación interactiva.

- **Cyber Academy (Chibitech):** Aplicación de escritorio que enseña ciberseguridad mediante escenarios simulados y aprendizaje adaptativo.

- **Piano Colors:** Aplicación educativa que usa tecnología interactiva para enseñar música a niños con síndrome de Down (Información externa proporcionada por el usuario).

5. Ecosistema Inmersivo (VR / AR)

Tecnologías que utilizan mundos virtuales o realidad aumentada.

- **ARmadillo:** Aplicación móvil de realidad aumentada para mejorar la comprensión lectora en niños de educación básica.
- **A darle átomos:** Laboratorio virtual de química en AR que utiliza reconocimiento de manos para experimentos seguros.
- **SimuRisk:** Simulador de realidad virtual para la formación inmersiva en prevención de riesgos laborales.

6. Ecosistema de Monitoreo (IoT / Sensores)

Sistemas que recopilan y visualizan datos de redes de sensores distribuidos.

- **AgroPred:** Sistema de agricultura de precisión que monitorea suelos y clima mediante sensores para optimizar cultivos de paltos.
- **CanarIA:** Sistema de monitoreo en tiempo real que utiliza IA para la detección de polvo en entornos industriales.
- **GerioTracker:** Sistema de monitoreo activo para la posición y seguimiento de adultos mayores en hogares de acogida.

Las siguientes clasificaciones corresponden a categorías que, si bien no se evidenciaron en la muestra específica de 2024, fueron incorporadas mediante un análisis crítico prospectivo. Esta inclusión se fundamenta en el conocimiento empírico adquirido por la autora durante su trayectoria académica y su posterior rol como ayudante de las asignaturas involucradas. Dicha experiencia en terreno permitió identificar vacíos metodológicos previos y anticipar la integración de tecnologías emergentes que poseen una alta probabilidad de figurar en futuras ediciones de la Feria de Software.

7. Ecosistema de Micro-interacciones (Wearables)

Software de interacción rápida para dispositivos corporales, con soluciones de visualización adaptadas a pantallas reducidas y uso de sensores biológicos. La pertinencia de esta categoría se valida empíricamente a través de la observación de proyectos previos al protocolo, como el caso de *Nura* (Feria de Software 2025), una aplicación diseñada para el seguimiento continuo de la frecuencia cardíaca y el movimiento mediante el uso del acelerómetro y giroscopio integrados en un reloj inteligente.

Qué incluye: Aplicaciones para smartwatches, monitores de salud y rastreadores de actividad física.

8. Ecosistema de Control y Robótica (Hardware Externo)

Software que actúa como interfaz lógica para operar, monitorear o automatizar máquinas y objetos físicos remotos. Aunque no se identificó un proyecto de esta naturaleza en la cohorte de 2024, su inclusión se justifica como una medida de preparación metodológica ante el avance de la industria (análisis prospectivo). La integración de interfaces que manejan hardware externo representa un desafío creciente en la Ingeniería de Software, por lo que el protocolo debe ser capaz de evaluar la fricción cognitiva en escenarios donde un error de usabilidad puede resultar en un daño físico al equipo operado.

Qué incluye: Paneles de control para drones, brazos robóticos, vehículos de exploración y software de domótica industrial o hardware externo en general.

3.2. Paso 2: Validación de la categorización preliminar

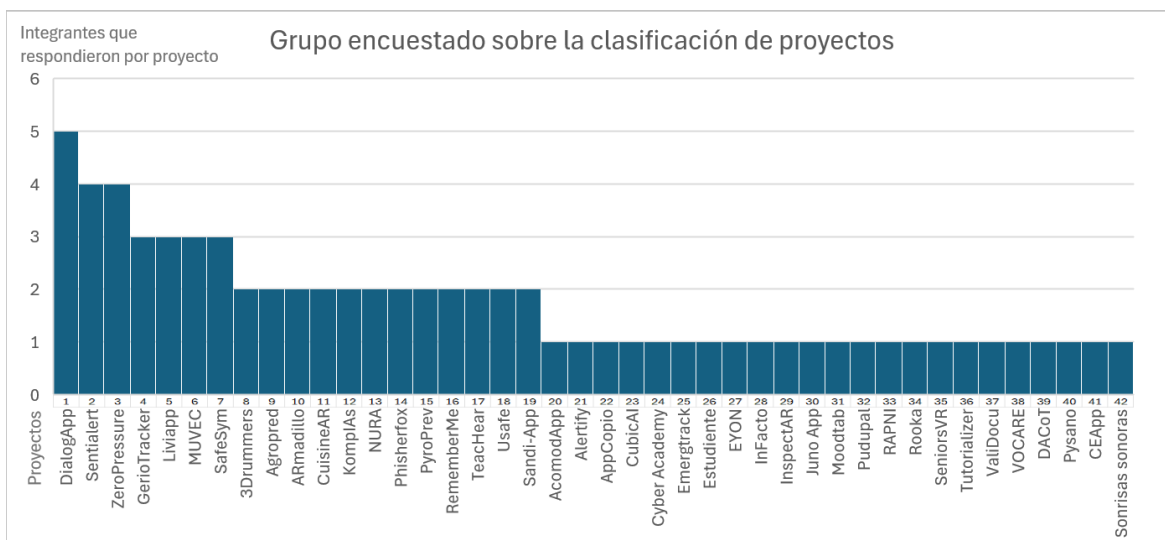
Posterior a la clasificación inicial se llevó a cabo una encuesta dirigida a los estudiantes que presentaron sus proyectos en las ediciones de la Feria de Software entre los años 2019 y 2025. La encuesta estructuró la recolección de información en tres etapas: **primero**, la caracterización del proyecto de software al que pertenecieron (Describe con detalle tu proyecto); **segundo**, la validación de su pertenencia a los ecosistemas tecnológicos establecidos (¿Que categoría incluirías basándote en la experiencia con el ramo?); y **tercero**, la identificación de

discrepancias(¿Sientes que falta alguna categoría?, si es si explica). Para los proyectos que no hallaron una correspondencia directa, se recolectaron comentarios explicativos y sugerencias sobre posibles nuevas categorías o mejoras al sistema de clasificación.

El propósito de este instrumento fue validar la pertinencia de las categorías propuestas, aprovechando la experiencia técnica y el conocimiento práctico de quienes han completado exitosamente el ciclo de las asignaturas

Como resultado del levantamiento de información, se obtuvieron 72 respuestas válidas correspondientes a participantes de 42 proyectos distintos, cuya distribución se detalla en el siguiente gráfico.

Figura 3.1: Cantidad de respuestas de distintos miembros por proyecto



Los siguientes gráficos son parte de la información recolectada a través de la encuesta, en el primero se puede ver que respondieron alumnos de Campus Casa Central y Campus San Joaquín junto a una única respuesta de Sede José Miguel Carrera(sede en que se empezó a realizar una feria de software en el año 2024), y en el segundo gráfico se puede observar el año al que pertenecieron los proyectos de los alumnos que respondieron, los cuales se encuentran en el rango de 2019 a 2025 a excepción de respuestas de alumnos del evento 2021(evento realizado de forma virtual debido a la pandemia).Esta diversidad de datos permite un análisis exhaustivo de la pertenencia de cada iniciativa a los ecosistemas preliminares.

Figura 3.2: Sede o Campus al que pertenecen quienes presentaron su proyecto

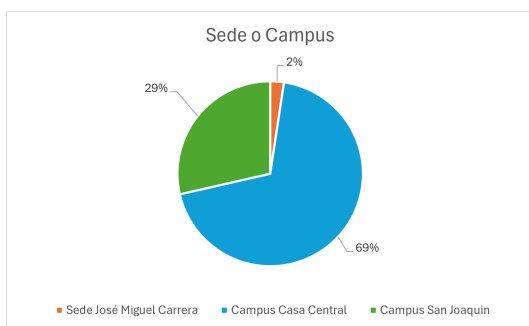
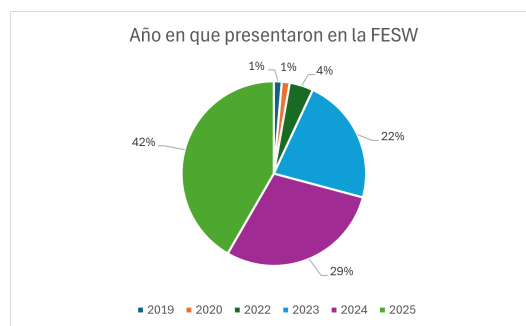


Figura 3.3: Respuestas por año de presentación del proyecto



Con el propósito de asegurar la validez metodológica del instrumento y obtener datos precisos para la reestructuración de los ecosistemas tecnológicos, la encuesta fue diseñada mediante un cuestionario semiestructurado. A los encuestados se les presentó inicialmente un listado detallado con las categorías preliminares definidas en la etapa anterior, para luego solicitarles que respondieran a las siguientes interrogantes, cada una formulada con un objetivo analítico específico:

- **Proyecto del que fuiste parte:** Su objetivo principal fue identificar unívocamente la solución de software desarrollada. Esto permitió realizar un cruce de datos con los registros históricos de la Feria de Software, validando la participación empírica del encuestado y el contexto temporal del proyecto.

- **Describe con detalle tu proyecto:** El propósito de esta pregunta fue recopilar una descripción funcional y técnica desde la perspectiva de los propios creadores. Al comprender las mecánicas centrales, la arquitectura lógica y las tecnologías subyacentes, se facilitó el mapeo y la validación de cada solución respecto a los ecosistemas preliminares propuestos.
- **¿Qué categoría incluirías basándote en la experiencia con el ramo?:** Esta interrogante buscó que los participantes realizaran una auto-clasificación de su proyecto dentro de la taxonomía proporcionada. Su objetivo fue medir el nivel de alineación y coherencia entre la categorización teórica preliminar y la percepción práctica de los estudiantes sobre la naturaleza de su propio desarrollo.
- **¿Sientes que falta alguna categoría? (Si la respuesta es sí, explica):** Constituyó el núcleo exploratorio del instrumento. Tras revisar el listado de categorías preliminares, se instó a los encuestados a realizar un análisis crítico para identificar vacíos tecnológicos o funcionales en la propuesta. Su objetivo directo fue levantar necesidades de clasificación no contempladas originalmente, proporcionando la justificación cualitativa para integrar ecosistemas emergentes y críticos.
- **Comentario o Sugerencia:** Su finalidad fue recolectar retroalimentación cualitativa abierta. Este espacio permitió a los participantes aportar observaciones metodológicas adicionales sobre el proceso de desarrollo de software, la experiencia técnica en las asignaturas o sugerencias estructurales que enriquecieran la formulación definitiva del protocolo de evaluación de UX.

3.2.1. Análisis de Hallazgos de la Encuesta

Para una correcta interpretación de los hallazgos obtenidos mediante el levantamiento empírico de información, resulta pertinente retomar la conceptualización metodológica establecida en la Sección 3.1.1. El análisis de las respuestas proporcionadas por los ex-participantes de la Feria de Software se fundamenta en la premisa de que el término **ecosistema** no se restringe de manera exclusiva al soporte físico o hardware utilizado por los proyectos. Por el

contrario, se comprende como un entorno integral de desarrollo que condiciona directamente la interacción del usuario final. Bajo esta perspectiva teórica, la evaluación de los datos recolectados se estructuró en torno a las dos dimensiones fundamentales que articulan la taxonomía propuesta: la **dimensión de soporte físico** y la **dimensión de arquitectura lógica y contexto operativo**. Esta aproximación multidimensional resulta crítica, ya que permitió analizar las sugerencias y discrepancias de los encuestados bajo un marco de trabajo estandarizado.

Con base en el análisis exhaustivo de las 72 respuestas recolectadas en la encuesta dirigida a participantes de ediciones anteriores de la Feria de Software, se identificaron hallazgos determinantes. La información clave que permitió actualizar y robustecer la clasificación preliminar de categorías se resume en los siguientes puntos:

3.2.1.1. Detección de Vacíos Tecnológicos y Funcionales

La retroalimentación cualitativa de los participantes reveló áreas técnicas y operativas que no estaban representadas adecuadamente en la estructura original. La justificación metodológica para la incorporación de estas nuevas dimensiones radica en sus paradigmas únicos de interacción:

- **Procesamiento de Audio:** Se identificó la necesidad de una categoría independiente para el software cuyo núcleo operativo reside en la captura y análisis de sonido. Desde la perspectiva de la Interacción Humano-Computadora, el audio exige un paradigma de evaluación distinto a la visión computacional. La usabilidad en este entorno queda supeditada a la tolerancia al ruido ambiental, la latencia de procesamiento y la retroalimentación sin apoyos gráficos tradicionales, requiriendo métricas orientadas estrictamente a la percepción sensorial.
- **Ciberseguridad:** Se evidenció que las plataformas orientadas a la seguridad informática poseen una complejidad estructural que justifica su tratamiento como un ecosistema independiente. La Experiencia de Usuario en este contexto trasciende la disposición visual (UI) para centrarse en la mitigación de la carga cognitiva y el manejo del estrés. En un panel de seguridad, el factor crítico de interacción es la confianza algorítmica del

operador frente a los datos y la prevención de errores bajo escenarios de alta presión, dimensiones que las heurísticas web tradicionales no logran capturar.

- **Investigación y Algoritmos:** Los encuestados sugirieron la inclusión de proyectos focalizados netamente en la innovación teórica (ej. optimización de algoritmos de compresión o escalamiento). La justificación de este vacío radica en que su *usuario final* no es el consumidor comercial masivo, sino perfiles altamente técnicos o integraciones de sistema (API). Por consiguiente, su evaluación UX debe priorizar la arquitectura de la información técnica, la legibilidad lógica y la eficiencia operativa por sobre el atractivo estético.
- **Sistemas Críticos en Tiempo Real:** Se destacó la omisión de arquitecturas lógicas donde la latencia mínima y la inmediatez de respuesta constituyen el factor central de éxito. En estos entornos, la UX no se audita por la satisfacción hedónica, sino por la fiabilidad transaccional y la prevención del error bajo presión temporal extrema. Un fallo de usabilidad o un cuello de botella visual en la interfaz de un sistema crítico compromete directamente la operación, exigiendo un marco de evaluación enfocado en la fatiga y la carga mental.

3.2.2. Refinamiento de Ecosistemas Existentes

La retroalimentación empírica obtenida no solo evidenció la omisión de ciertas áreas tecnológicas, sino que expuso la necesidad metodológica de refinar y delimitar los ecosistemas preliminares. El propósito central de esta reclasificación es abandonar el enfoque tradicional que agrupa el software exclusivamente por su soporte físico (ej. plataformas web o aplicaciones móviles). En su lugar, el objetivo (*para qué se clasifica*) es estructurar una taxonomía operativa que permita vincular de manera unívoca cada ecosistema con los instrumentos psicométricos y heurísticos especializados que se definirán en las etapas posteriores del protocolo.

Para lograr esta precisión, el proceso de refinamiento se ejecutó estableciendo nuevas características de agrupación (*cómo se clasifica*). Los proyectos dejaron de categorizarse por su aspecto visual para ser clasificados según su **paradigma de interacción**, el **nivel de carga**

cognitiva exigida al operador y el **factor crítico de éxito** del sistema (ej. retención pedagógica versus confianza algorítmica).

Bajo estos nuevos criterios de clasificación, las categorías originales evolucionaron para reflejar su verdadera arquitectura lógica:

- **De Entretenimiento a Formación Interactiva (EdTech):** Se detectó que categorizar genéricamente como “Videojuegos” resultaba metodológicamente excluyente para las herramientas de *e-learning* y simuladores pedagógicos. En la Interacción Humano-Computadora, un videojuego tradicional busca puramente el ensimismamiento, disfrute y gratificación personal del usuario. Por el contrario, la principal característica del software de formación interactiva es su arquitectura dual: debe mantener el enganche lúdico sin que este enmascare u obstaculice la carga cognitiva necesaria para cumplir un objetivo de aprendizaje formal.
- **Evolución del Monitoreo a Predictivo (IoT / Machine Learning):** La retroalimentación evidenció que los sistemas de monitoreo contemporáneos han transitado desde la simple visualización pasiva de datos hacia el análisis predictivo. Al integrar modelos de Inteligencia Artificial para inferir comportamientos futuros, la característica central de la Experiencia de Usuario cambia drásticamente. La evaluación ya no depende solo de la legibilidad gráfica del *dashboard*, sino de medir empíricamente si el usuario desarrolla una confianza óptima para delegar la toma de decisiones en el sistema automatizado.
- **Amplitud en Ecosistemas de Gestión (SaaS y Big Data):** Se concluyó que la categoría original de “Gestión” carecía de la granularidad necesaria. La característica definitoria de los proyectos orientados a Software as a Service (SaaS) y plataformas de ciencia de datos es que exigen al usuario el procesamiento de grandes volúmenes de información (alta densidad de datos). Por consiguiente, su evaluación UX debe priorizar la eficiencia transaccional, la prevención de errores en flujos de trabajo prolongados y la mitigación de la fatiga visual en pantallas de escritorio.

Esta síntesis cualitativa fundamenta la necesidad de reestructurar el modelo preliminar. Al integrar el procesamiento sensorial, formalizar los entornos críticos y refinar los ecosistemas

predictivos y educativos, se sientan las bases para consolidar la taxonomía definitiva que se detalla a continuación en la Sección 3.3 (Paso 3). Esta reestructuración es el eslabón fundamental que permitirá asignar las métricas exactas a la base técnica de cada proyecto.

3.3. Paso 3: Reestructuración de las categorías

Basado en las definiciones iniciales y las sugerencias clave de los participantes de la Feria de Software de años anteriores, se han reestructurado las categorías para integrar las propuestas más relevantes, como la creación de un ecosistema de audio, la separación de software educativo de los videojuegos y la formalización de la ciberseguridad.

1. Ecosistema de Gestión, SaaS y Big Data (Web / Desktop)

Se amplía para incluir plataformas de análisis y ciencia de datos. Está orientado a la productividad y al procesamiento de grandes volúmenes de información.

Priorización de UX: Se debe priorizar la eficiencia en la tarea y la densidad de información. Se recomiendan Evaluaciones Heurísticas centradas en la visibilidad del estado del sistema y la prevención de errores en flujos transaccionales complejos.

2. Ecosistema de Movilidad y Servicios Transaccionales (Mobile)

Aplicaciones para uso en movimiento que no dependen de sensores complejos de imagen o audio.

Priorización de UX: El foco debe estar en la usabilidad en contextos de distracción y la “zona del pulgar”. Se deben medir los tiempos de carga y la facilidad de navegación con una sola mano mediante pruebas de guerrilla con usuarios reales.

3. Ecosistema de Percepción Sensorial (Captura de Visión y Audio)

Siguiendo las sugerencias de los participantes, esta categoría ahora integra tanto la visión computacional como el procesamiento de audio. Incluye proyectos que usan el micrófono o la cámara como núcleo operativo.

Priorización de UX: La métrica crítica es la latencia de respuesta y la retroalimentación de éxito de captura. La UX se evalúa midiendo la tolerancia al ruido ambiental

(lumínico o sonoro).

4. **Ecosistema de Formación Interactiva (Educación y EdTech)**

Se reformula la antigua categoría de “Videojuegos” para incluir software educativo interactivo que no necesariamente constituye un juego, respondiendo a la necesidad de clasificar rigurosamente las herramientas pedagógicas.

Priorización de UX: Debe evaluarse mediante la carga cognitiva y la retención de aprendizaje. Se utilizan encuestas de satisfacción post-uso y métricas de completitud de objetivos pedagógicos.

5. **Ecosistema Inmersivo y Háptico (VR / AR)**

Mantiene el uso de mundos virtuales y realidad aumentada, integrando además el reconocimiento de gestos.

Priorización de UX: La evaluación debe centrarse primordialmente en la ergonomía y la prevención de la cinetosis (*motion sickness*). Se recomienda el uso de escalas de “Presencia” y “Confort” para asegurar que la interacción espacial no fatigue al usuario.

6. **Ecosistema de Ciberseguridad y Sistemas Críticos**

Nueva categoría consolidada a partir de los vacíos detectados en la fase de validación empírica. Se justifica su inclusión independiente bajo la dimensión de *arquitectura lógica y contexto operativo*. En este ecosistema, la complejidad no recae en el uso de un hardware físico no convencional, sino en el volumen de amenazas, la carga cognitiva requerida para el monitoreo y la inmediatez de respuesta vital para la operación.

Priorización de UX: El eje central de medición abandona la usabilidad funcional básica para centrarse en la percepción de confianza algorítmica y la carga mental. Se evalúa la claridad visual de los indicadores de estado (seguro/vulnerable) y el estrés u operabilidad bajo presión frente a una amenaza detectada por el sistema.

7. **Ecosistema de Monitoreo Predictivo e IoT (Sensores)**

Evoluciona desde la simple recopilación pasiva de datos hacia el análisis predictivo mediante modelos de *Machine Learning*.

Priorización de UX: Se debe evaluar la claridad de la visualización de datos (*Data Viz*). Es crucial medir si el usuario puede interpretar correctamente las tendencias y

predicciones futuras sin necesidad de ser un experto en analítica de datos.

8. Ecosistema de Wearables y Salud Digital

Especializado en dispositivos corporales y microinteracciones, con un fuerte enfoque en la accesibilidad y el monitoreo continuo de parámetros de salud.

Priorización de UX: El foco es la intrusividad mínima y la lectura rápida (*glancability*). Se evalúa qué tan rápido el usuario o el cuidador puede obtener información crítica sin interrumpir su actividad principal.

Con los hallazgos de la encuesta se puede ver que se logró realizar la reestructuración de los ecosistemas tecnológicos integrando el procesamiento de audio, la ciberseguridad, los sistemas de tiempo real y la separación de herramientas pedagógicas serias de los videojuegos tradicionales. Este refinamiento delimita las fronteras operativas y funcionales de cada software para responder de manera precisa a la diversidad tecnológica de los proyectos, obteniendo la clasificación oficial y que será utilizada finalmente en la creación del protocolo.

3.4. Paso 4: Asignación estratégica de metodologías y métricas de UX

En esta fase se realizó una revisión exhaustiva de las metodologías e instrumentos de experiencia de usuario disponibles en la literatura técnica y científica, utilizando la información crítica recopilada desde el estado del arte. El propósito de este análisis bibliográfico fue contrastar las herramientas psicométricas validadas con los ocho ecosistemas tecnológicos reestructurados en el paso anterior, determinando la pertinencia, sensibilidad y viabilidad de cada escala según la naturaleza de los proyectos.

Como resultado de este cruce analítico, se estructuró el marco metodológico central de esta propuesta. Dicho marco se materializa como una **matriz operativa de evaluación** que supera el enfoque tradicional homogéneo donde se aplica una misma encuesta genérica a cualquier software para estandarizar rutas de auditoría específicas. En términos prácticos,

este marco metodológico vincula de manera unívoca cada ecosistema con una dupla de instrumentos psicométricos y heurísticos especializados (metodología de triangulación). Esta arquitectura de asignación es lo que permite medir empíricamente los requerimientos particulares de cada dominio, tales como la fricción táctil en el rendimiento móvil, la gestión de la carga mental en sistemas críticos o el impacto ergonómico en dispositivos corporales.

3.4.1. Criterios de Selección y Modelo de Triangulación Metodológica

Para garantizar que la asignación de escalas cuente con un sustento científico riguroso, la toma de decisiones no se realizó de manera aislada. A partir de los hallazgos del estado del arte, se determinó que la evaluación de tecnologías heterogéneas requería un modelo de triangulación metodológica intramétodo (Within-Method Triangulation) [13].

Este modelo asocia estratégicamente dos instrumentos de medición por categoría para capturar de forma complementaria tanto la usabilidad de la interfaz como los factores críticos del entorno operativo. Para justificar esta arquitectura de evaluación, la matriz se estructuró sobre tres ejes fundamentales, cada uno con un objetivo analítico claro:

1. El Eje de Granularidad Temporal: Micro-Táctico vs. Macro-Estratégico **Objetivo metodológico del eje:** Contrastar la fricción cognitiva inmediata generada en una acción puntual (micro) frente a la percepción retrospectiva de todo el sistema (macro), evitando así que el sesgo de recencia o la memoria a corto plazo del usuario distorsionen el diagnóstico global.

Bajo este objetivo, se estructuró la triangulación de los ecosistemas basados en la eficiencia transaccional pura:

- **Ecosistema 1 (Gestión y SaaS) y Ecosistema 2 (Movilidad):** Se vincula la métrica táctica SEQ [23](evaluada inmediatamente tras una tarea) con cuestionarios holísticos post-test como el SUS [14] o el SUPR-Qm [16]. De este modo, si el puntaje global macro es deficiente, el equipo desarrollador puede rastrear con precisión quirúrgica,

mediante los picos de baja puntuación en la SEQ, qué pantalla o flujo interactivo específico rompió la percepción del sistema.

2. El Eje de Especialización: Usabilidad de Software vs. Factores Críticos de Frontera

Objetivo metodológico del eje: Complementar una métrica estándar de usabilidad con una escala psicológica o ergonómica especializada, aplicable cuando la base tecnológica introduce variables de estrés ajenas al software tradicional (como el confort físico, la confianza algorítmica o la motivación lúdica).

Este eje rige la triangulación en los ecosistemas donde la interfaz gráfica está supeditada a factores de frontera:

- **Ecosistema 4 (Formación Interactiva):** Se cruza UMUX-LITE [26] con GUESS-18 [17]. El objetivo es aislar la atribución lúdica (medida por GUESS) de la eficiencia operativa (medida por UMUX), evitando el fenómeno de enmascaramiento donde un software entretenido oculta mecánicas de aprendizaje ineficientes.
- **Ecosistema 6 (Ciberseguridad y Sistemas Críticos):** Se enlaza Jian Trust [22] con NASA-TLX [24]. En sistemas críticos, la usabilidad visual pasa a segundo plano; el objetivo es diagnosticar si el error del operador es causado por una crisis de confianza hacia la Inteligencia Artificial (Jian Trust) o por una sobrecarga de estrés y fatiga bajo presión (NASA-TLX).
- **Ecosistemas 7 (Monitoreo Predictivo) y 8 (Wearables):** Se cruza la fricción táctica (SEQ [23]) con la Escala Jian Trust [22] y la CRS [20], respectivamente. El objetivo es validar que el éxito de la interacción no dependa solo de los botones en pantalla, sino de la confianza en los datos proyectados (IoT) o del confort biomecánico del dispositivo en el cuerpo (Wearables).

3. El Eje de Validación por Terceros e Inmersión Simétrica

Objetivo metodológico del eje: Correlacionar la percepción subjetiva del usuario con restricciones fisiológicas o ambientales objetivas en entornos cerrados, donde la opinión declarativa tradicional es insuficiente para diagnosticar un fallo de hardware.

Este eje articula los ecosistemas altamente dependientes de sensores físicos:

- **Ecosistema 3 (Percepción Sensorial):** Se conecta el método analítico experto (PURE [25]) con la percepción del usuario (SEQ [23]). El objetivo es discriminar si una mala calificación de usabilidad se debe a un flujo de diseño interactivo deficiente o a un fallo ambiental que el micrófono/cámara no supo procesar.
- **Ecosistema 5 (VR/AR):** La conexión entre el PQ (Presencia)[18] y el SSQ (Cinetosis)[19] responde a un objetivo biológico: verificar empíricamente que el nivel de inmersión cognitiva en el mundo virtual (PQ) no esté sabotando el sistema vestibular del usuario y generando mareos por latencia de renderizado (SSQ).

3.4.2. Clasificación de Ecosistemas Tecnológicos y Matriz de Escalas UX

Con el propósito de adaptar la evaluación de la experiencia de usuario a las características particulares de cada tipo de tecnología, y una vez establecida la lógica científica de selección y triangulación derivada del estado del arte, a continuación se presenta la estructuración definitiva de los ocho ecosistemas tecnológicos identificados, detallando los instrumentos asignados, sus rangos operativos, enfoques de medición y la justificación académica que valida su implementación conjunta.

3.4.2.1. Asignación de escalas de evaluación según ecosistema tecnológico

Con base en el análisis del estado del arte y las características particulares de cada dominio de aplicación, se definió una asignación específica de escalas e instrumentos de evaluación UX para cada ecosistema tecnológico identificado. Esta clasificación busca garantizar que las métricas seleccionadas sean coherentes con las necesidades de interacción, los riesgos operacionales y los objetivos de experiencia de usuario de cada contexto.

Ecosistema 1: Gestión, SaaS y Big Data (Web/Desktop)

Escala asignada: SUS (System Usability Scale)[14] + %SEQ (Single Ease Question)[23].

Rango original: SUS: 0 a 5 puntos | SEQ: 1 a 7 puntos..

Descripción del tipo de desarrollos: Plataformas orientadas a entornos de oficina o escritorio que manejan flujos administrativos y operativos complejos. Incluye sistemas de planificación de recursos empresariales (ERP), herramientas de software como servicio (SaaS), tableros de control analítico de datos (dashboards), utilitarios de productividad laboral y aplicaciones web dedicadas a la manipulación masiva de registros. En este ecosistema, la interacción transaccional se basa en el intercambio continuo y detallado de información (creación, lectura, actualización y borrado de datos complejos), independientemente de si involucra o no cobros monetarios.

Enfoque de medición: Evaluación de la usabilidad retrospectiva global del sistema combinada con la medición de la fricción inmediata en flujos críticos específicos de alta densidad de datos.

Justificación: Las plataformas de gestión y SaaS operan bajo una *arquitectura lógica* de alta densidad de datos, donde el *contexto operativo* exige mantener la atención sostenida y gestionar flujos transaccionales complejos. En este entorno, la escala *SUS* proporciona un diagnóstico macro-estratégico sobre la facilidad de aprendizaje y eficiencia global del sistema. Sin embargo, para evitar que la evaluación global diluya problemas específicos de la arquitectura de información, se aplica la métrica micro-táctica *SEQ* inmediatamente después de tareas críticas. Esta triangulación permite identificar si la fricción cognitiva se genera por la sobrecarga del modelo mental en un paso transaccional específico, aislándolo de la percepción general del software.

Ecosistema 2: Movilidad y Servicios Transaccionales (Mobile)

Escala asignada: SEQ (Single Ease Question)[23] + SUPR-Qm (Standardized User Experience Percentile Rank Questionnaire - Mobile / Short Form) [16].

Rango original: SEQ: 1 a 7 puntos | SUPR-Qm: Ítems 1 al 12 en escala Likert de 1 a 5 puntos; Ítem 13 (NPS) en escala de 0 a 10 puntos.

Descripción del tipo de desarrollos: Aplicaciones móviles nativas o híbridas para teléfonos inteligentes y tabletas, diseñadas para apoyar actividades operativas en contextos de movilidad. Comprende plataformas de comercio electrónico, servicios logísticos y de delivery, redes sociales, mensajería, geolocalización y herramientas de productividad utilizadas en terreno. El enfoque transaccional de esta categoría refiere a interacciones ágiles y secuenciales (ej. reportar un incidente en calle, confirmar una entrega, o enviar un formulario de fiscalización) que deben ejecutarse de forma rápida bajo contextos de atención fragmentada.

Enfoque de medición: Balance entre la facilidad de operación táctica en movimiento (medición post-tarea inmediata) y la evaluación macro-holística de la experiencia móvil (medición post-test declarativa), abarcando usabilidad, confianza transaccional, atractivo visual e intención de recomendación (NPS).

Justificación: El entorno móvil está fuertemente condicionado por su *soporte físico*, imponiendo restricciones de pantalla y exigiendo interacciones táctiles rápidas bajo contextos operativos de atención fragmentada. Para evaluar esta dimensión física, la métrica *SEQ* se aplica post-tarea, midiendo la fricción táctica y visual inmediata en flujos como la zona del pulgar. Simultáneamente, la *arquitectura lógica* móvil exige el procesamiento de transacciones seguras; por ello, se complementa con la métrica post-test *SUPR-Qm*, la cual audita de forma holística cuatro factores críticos: usabilidad, apariencia, intención de lealtad (NPS) y, fundamentalmente, la confianza en el manejo de datos, asegurando que las restricciones del hardware no comprometan la credibilidad del servicio.

Ecosistema 3: Percepción Sensorial (Captura de Visión y Audio)

Escala asignada: Evaluación PURE (Practical Usability Rating by Experts)[25] + SEQ (Single Ease Question)[23].

Rango original: PURE: Puntuación acumulada de fricción (1 a 3 por paso) | SEQ: 1 a 7 puntos.

Descripción del tipo de desarrollos: Sistemas y aplicaciones basados en software y hardware cuyo núcleo funcional depende del procesamiento en tiempo real de señales ambientales y biológicas capturadas a través de cámaras, micrófonos u otros sensores. Esta categoría abarca algoritmos de visión artificial y computacional aplicados a la clasificación de objetos, el

control de acceso por reconocimiento facial y la detección de fatiga, así como herramientas de reconocimiento, transcripción o análisis automatizado de voz. Asimismo, incluye sistemas de traducción en tiempo real de lenguaje de señas, flujos interactivos guiados por gestos, autenticación biométrica, utilitarios para el filtrado adaptativo de señales y herramientas de asistencia tecnológica diseñadas para personas con discapacidades sensoriales.

Enfoque de medición: Cuantificación técnica y heurística, realizada por un evaluador en rol de experto UX (docentes, ayudantes o auditores del equipo), sobre la latencia de respuesta, precisión de captura y tolerancia al ruido ambiental.

Justificación: Los sistemas de percepción sensorial dependen intrínsecamente del *soporte físico* periférico (cámaras, micrófonos) y están expuestos a variables ambientales incontrolables (ruido, iluminación) que afectan su *arquitectura lógica* de procesamiento. Debido a que los usuarios finales rara vez pueden diagnosticar si un fallo es de interacción gráfica o del sensor, se aplica el método experto *PURE* para auditar de forma heurística la carga cognitiva y el esfuerzo físico de cada paso del flujo de captura. Esta evaluación analítica y predictiva de la interfaz se triangula con la escala *SEQ* aplicada al usuario empírico, permitiendo contrastar de manera objetiva si la fricción teórica calculada por los expertos coincide con la resistencia táctica real experimentada por el sujeto al interactuar con el hardware.

Ecosistema 4: Formación Interactiva (Educación y EdTech)

Escala asignada: Escala GUESS-18 (Game User Experience Satisfaction Scale)[17] + UMUX-LITE [26].

Rango original: GUESS-18: 1 a 7 puntos por ítem | UMUX-LITE: 1 a 7 puntos.

Descripción del tipo de desarrollos: Soluciones de software educativo interactivo y orientadas a los procesos de enseñanza y aprendizaje a través del enganche interactivo. Incluye plataformas de e-learning adaptativo, plataformas de gestión del aprendizaje (LMS) gamificadas, videojuegos serios (serious games) destinados a la enseñanza o diseñados explícitamente para la capacitación técnica corporativa o escolar, simuladores académicos virtuales de bajo costo y de laboratorios académicos, aplicaciones educativas móviles para el desarrollo de habilidades específicas (idiomas, ciencias) o herramientas gamificadas de entrenamiento

profesional que combinan el cumplimiento de objetivos pedagógicos con dinámicas recreativas y de enganche.

Enfoque de medición: Evaluación de absorción, disfrute, usabilidad y gratificación personal en experiencias educativas gamificadas. Evaluación simultánea de los componentes de enganche lúdico/pedagógico y los atributos puros de utilidad del sistema computacional educativo.

Justificación: Los entornos de formación interactiva operan bajo una *arquitectura lógica* dual, donde el cumplimiento de objetivos pedagógicos convive con la necesidad de mantener el compromiso del usuario. En este contexto, la escala *GUESS-18* permite medir las dimensiones motivacionales y lúdicas (como el ensimismamiento y la narrativa). Sin embargo, para evitar el fenómeno de enmascaramiento donde una alta motivación oculta deficiencias operativas o viceversa, se integra de forma complementaria la escala *UMUX-LITE*. Esta triangulación permite discriminar empíricamente si una fricción en el aprendizaje es causada por un diseño pedagógico ineficaz o por una barrera en la usabilidad pura del sistema de software.

Ecosistema 5: Ecosistema Inmersivo y Háptico (VR/AR)

Escala asignada: PQ (Presence Questionnaire)[18] + SSQ (Simulator Sickness Questionnaire)[19].

Rango original:PQ: 1 a 7 puntos | SSQ: Puntuación ponderada de síntomas de cinetosis.

Descripción del tipo de desarrollos: Aplicaciones espaciales desarrolladas para cascos o gafas de realidad virtual (VR), experiencias móviles basadas en la superposición de capas interactivas mediante realidad aumentada (AR), sistemas interactivos guiados por reconocimiento de gestos tridimensionales o simuladores de prevención de riesgos que interactúan con hardware de retroalimentación física.

Enfoque de medición: Correlación entre la inmersión psicológica ("sentimiento de estar allí") y los efectos adversos de la fatiga física o desorientación espacial provocados por el hardware.

Justificación: Las tecnologías inmersivas están fuertemente condicionadas por la dimensión del *soporte físico*, donde el hardware (visores y sensores espaciales) interactúa de manera directa con el sistema vestibular y perceptivo del individuo. La evaluación exige medir la

inmersión psicológica a través del Cuestionario de Presencia (*PQ*), cuantificando el grado de naturalidad y control en el entorno virtual. Simultáneamente, resulta crítico auditar las restricciones fisiológicas impuestas por el dispositivo; para ello, el *SSQ* diagnostica la cinetosis aislando factores oculomotores, náuseas y desorientación espacial, garantizando que el diseño de locomoción y renderizado no genere un impacto perjudicial en la salud del usuario.

Ecosistema 6: Ciberseguridad y Sistemas Críticos

Escala asignada: Escala de Confianza de Jian (Jian Trust Scale)[22] + NASA-TLX (Task Load Index)[24].

Rango original: Jian Trust: 1 a 7 puntos | NASA-TLX: 0 a 100 puntos (índice global ponderado de carga de trabajo).

Descripción del tipo de desarrollos: Consolas de comando para auditoría de código, herramientas de detección activa de intrusiones, paneles de control para la respuesta ante incidentes informáticos, software de cifrado de comunicaciones, sistemas de respuesta automatizada ante vulnerabilidades de red y entornos de monitorización industrial en tiempo real donde un retardo en las alertas compromete la seguridad.

Enfoque de medición: Evaluación integral de la confianza del operador en los algoritmos de seguridad combinada con la cuantificación de la carga de trabajo subjetiva (demanda mental, temporal, esfuerzo y frustración) impuesta por la interfaz crítica.

Justificación: Los entornos de ciberseguridad y sistemas de alta criticidad operan bajo la dimensión de arquitectura lógica, sometiendo al operador a escenarios de elevada presión temporal y toma de decisiones bajo incertidumbre, independientemente del hardware utilizado. En este contexto operativo, la *Escala de Confianza de Jian* permite medir empíricamente la relación de confianza entre el usuario y las automatizaciones del algoritmo (fidelidad, integridad y fiabilidad percibida). Por su parte, la inclusión del cuestionario *NASA-TLX* resulta fundamental para diagnosticar el desgaste cognitivo, identificando si la fricción en el centro de monitoreo es provocada por un diseño estresante, demandas mentales excesivas o frustración operativa durante la gestión de incidentes críticos.

Ecosistema 7: Monitoreo Predictivo e IoT

Escala asignada: Escala de Confianza de Jian (Jian Trust Scale)[22] + SEQ (Single Ease Question)[23].

Rango original: Jian Trust: 1 a 7 puntos | SEQ: 1 a 7 puntos

Descripción del tipo de desarrollos: Dashboards interactivos alimentados por redes de sensores físicos distribuidos, aplicaciones de agricultura de precisión basadas en telemetría de suelos, software de mantenimiento predictivo industrial mediante machine learning, sistemas domóticos de control residencial o visualizadores de analítica avanzada que traducen datos crudos de hardware remoto en tendencias legibles. Automatización hogareña o urbana (Smart Cities) y plataformas de mantenimiento predictivo industrial con proyecciones y visualización de anomalías basadas en Machine Learning.

Enfoque de medición: Medición de la confianza algorítmica en las proyecciones e inferencias de Machine Learning (post-test) combinada con la evaluación de la facilidad inmediata de extracción de insights e interacción con gráficos interactivos (post-tarea).

Justificación: Las plataformas de monitoreo predictivo e IoT operan bajo una *arquitectura lógica* avanzada, delegando el análisis de telemetría a modelos de Machine Learning y automatización. El éxito de este *contexto operativo* no depende del hardware tradicional de visualización, sino de la percepción de seguridad algorítmica; por ello, la *Escala Jian de Confianza* resulta imperativa para diagnosticar de manera empírica la fiabilidad, integridad y posible desconfianza activa hacia las alertas del algoritmo. Para garantizar que una baja calificación en confianza se deba efectivamente a un escepticismo hacia la Inteligencia Artificial y no a una interfaz gráfica incomprensible, se aplica la métrica *SEQ* tras la extracción de datos, aislando la fricción táctica de la visualización (Data Viz) del nivel de delegación exitosa en el sistema automatizado.

Ecosistema 8: Wearables y Salud Digital

Escala asignada: CRS (Comfort Rating Scale)[20] + SEQ (Single Ease Question)[23].

Rango original: CRS: 0 a 20 puntos por dimensión (evaluación multidimensional de 6

subescalas donde 0 representa confort absoluto y 20 molestia extrema)| SEQ: 1 a 7 puntos.

Descripción del tipo de desarrollos: Aplicaciones secundarias adaptadas a relojes inteligentes (smartwatches), software de monitorización remota de signos vitales para pacientes crónicos, interfaces de usuario para dispositivos médicos vestibles portátiles o complementos de software que procesan información continua de bandas de entrenamiento físico o sensores corporales integrados.

Enfoque de medición: Evaluación integral de la carga ergonómica, física y psicológica del hardware mediante análisis estadístico inferencial (ANOVA [12]), combinada con la medición post-tarea de la usabilidad y facilidad de interacción en pantallas de formato reducido.

Justificación: Los dispositivos vestibles representan la máxima integración del *soporte físico* con el cuerpo humano, exigiendo que la evaluación trascienda la pantalla. La escala CRS permite cuantificar la carga ergonómica, cinemática y psicológica impuesta por el hardware de manera multidimensional, requiriendo un análisis de varianza (ANOVA) para validar empíricamente su confort en estado dinámico. A nivel de *arquitectura lógica*, estos dispositivos operan mediante micro-interacciones de alta velocidad; por lo tanto, la aplicación post-tarea de la SEQ asegura que la interfaz gráfica permita una lectura y respuesta rápida (glancability) sin interrumpir la actividad biomecánica principal del usuario.

Tabla 3.1: Matriz de Síntesis: Ecosistemas, Escalas de Captura y Umbrales de Éxito UX

Ecosistema Tecnológico	Escala Asignada	Formato de Captura (Pre-conversión)	Rango Tabulado (Post-conversión)	Umbral Buena UX (Aceptable a Óptimo)	Umbral Mala UX (Deficiente a Crítico)
1: Gestión, SaaS y Big Data (Web/Desktop)	SUS	Likert 5 puntos (1-5)	0 a 100 puntos	68 a 100 puntos	Inferior a 68 puntos
	SEQ	Likert 7 puntos (1-7)	1 a 7 puntos	6 a 7 puntos	1 a 5 puntos
2: Movilidad y Servicios Transaccionales (Mobile)	SUPR-QM	Likert 5 pts (ítems 1-12) y 0-10 (ítem 13)	0 a 100 puntos	68.0 a 100.0 puntos	Inferior a 68.0
	SEQ	Likert 7 puntos (1-7)	1 a 7 puntos	6 a 7 puntos	1 a 5 puntos
3: Percepción Sensorial (Captura de Visión y Audio)	PURE	Juicio experto (1-3)	1 a 3 (por paso)	Promedio global ≤ 2.0 (ningún paso en 3)	Promedio global >2.0 (o cualquier paso en 3)
	SEQ	Likert 7 puntos (1-7)	1 a 7 puntos	6 a 7 puntos	1 a 5 puntos
4: Formación Interactiva (Educación y EdTech)	GUESS-18	Likert 7 puntos (1-7)	9 a 63 puntos	46 a 63 puntos	9 a 45 puntos
	UMUX-LITE	Likert 7 puntos (1-7)	0 a 100 puntos	68 a 100 puntos	Inferior a 68 puntos
5: Ecosistema Inmersivo y Háptico (VR/AR)	PQ	Likert 7 puntos (1-7)	32 a 224 puntos	160 a 224 puntos	32 a 159 puntos
	SSQ	Escala 4 puntos (0-3)	>0 puntos	0 a 20 puntos	Superior a 20 puntos
6: Ciberseguridad y Sistemas Críticos	Jian Trust	Likert 7 puntos (1-7)	12 a 84 puntos	60 a 84 puntos	12 a 59 puntos
	NASA-TLX	Pares de peso + Escala 0-100 (incrementos 5)	0 a 100 puntos	0.0 a 66.6 puntos	66.7 a 100.0 puntos

7: Monitoreo Predictivo e IoT	Jian Trust	Likert 7 puntos (1-7)	12 a 84 puntos	60 a 84 puntos	12 a 59 puntos
	SEQ	Likert 7 puntos (1-7)	1 a 7 puntos	6 a 7 puntos	1 a 5 puntos
8: Wearables y Salud Digital	CRS	Continua (0-20)	0 a 20 por dimensión	0 a 10 puntos (en todas las dimensiones)	11 a 20 puntos (en cualquier dimensión)
	SEQ	Likert 7 puntos (1-7)	1 a 7 puntos	6 a 7 puntos	1 a 5 puntos

La consolidación expuesta en la Tabla 3.1 trasciende la mera clasificación teórica; constituye el eje articulador que garantiza la viabilidad empírica de esta propuesta. Al delimitar umbrales cuantitativos de éxito y fracaso para cada dupla metodológica, se elimina la subjetividad inherente a las revisiones superficiales de interfaz, permitiendo diagnosticar desde la fricción transaccional hasta la fatiga ergonómica o la desconfianza algorítmica según corresponda. Esta matriz operativa asegura que la auditoría de Experiencia de Usuario se ajuste con rigor analítico a la arquitectura lógica y al soporte físico de cada proyecto, resguardando la validez de los resultados frente a la heterogeneidad tecnológica. Con la asignación estratégica de los instrumentos psicométricos finalizada, el desafío subsiguiente radica en integrar estas métricas de manera orgánica dentro del ciclo de desarrollo ágil de los estudiantes. Esto sienta las bases directas para la estructuración del plan de trazabilidad y el diseño del marco de evaluación práctico (UX-Scorecard), los cuales conforman el protocolo definitivo abordado en el siguiente paso.

3.5. Paso 5: Diseño del protocolo de evaluación UX para proyectos de la FESW

3.5.1. Paso 5: Diseño del protocolo de evaluación UX para proyectos de la FESW

El diseño del protocolo constituye la culminación práctica de la propuesta metodológica. Su objetivo es transformar el marco teórico y la matriz de triangulación en un instrumento estandarizado, replicable y auditable, que guíe a los equipos de desarrollo desde la fase de prototipado hasta el despliegue final de su proyecto en la Feria de Software.

Para salvaguardar el rigor analítico de la presente memoria y mantener la fluidez narrativa del documento principal, el compendio íntegro de instrucciones, fórmulas matemáticas de tabulación y rúbricas psicométricas ha sido consolidado en un documento independiente. Este documento, denominado **Protocolo de Evaluación UX para proyectos desarrollados para la Feria de software**, se encuentra disponible para su consulta en el **Anexo A** del presente trabajo, al cual se puede acceder de manera directa si se requiere profundizar en los lineamientos de ejecución de cada métrica.

La estructura y los componentes que rigen dicho protocolo anexo se definen a continuación:

- **Introducción y propósito del protocolo:** Sección orientada a contextualizar a los equipos de desarrollo sobre la importancia de la Experiencia de Usuario como pilar estratégico, estableciendo las reglas de uso del manual durante los *sprints* de las asignaturas.
- **Marco heurístico de usabilidad elegido:** Presentación de las 10 Heurísticas de Usabilidad de Jakob Nielsen [6], estableciéndolas como la línea base de inspección que todo proyecto debe cumplir previo a las pruebas empíricas con usuarios reales.
- **Presentación de los ecosistemas tecnológicos:** Descripción general de las ocho categorías tecnológicas consolidadas en el Paso 3 (3.3), delineando brevemente las características de soporte físico y arquitectura lógica de cada una.

- **Escalas de evaluación psicométrica:** Catálogo descriptivo de los instrumentos estandarizados que componen el protocolo (tales como SUS, SUPR-Qm, GUESS-18, PQ, SSQ, CRS, Jian Trust, NASA-TLX, UMUX-LITE, SEQ y PURE), indicando su propósito de medición teórico (ej. carga cognitiva, fricción táctica, inmersión, confort ergonómico), sin detallar aún su asignación.
- **Ruta Diagnóstica (Árbol de Decisión):** Instrumento que opera como el primer filtro del marco de evaluación. Guía a los equipos a través de interrogantes dicotómicas fundamentadas en las dimensiones de hardware y contexto operativo, permitiéndoles identificar unívocamente a qué ecosistema tecnológico pertenece su solución de software para evitar sesgos en la selección de métricas.
- **Explicación detallada por ecosistema y guías operativas:** Desglose operativo donde se asocia la dupla de escalas correspondientes a cada ecosistema (Triangulación Metodológica). Se explica el caso de uso específico para cada entorno tecnológico. Asimismo, se hace referencia a las guías de uso detalladas, fórmulas de tabulación (como regresiones o inversiones matemáticas) y reglas de aplicación temporal, las cuales se adjuntan en su totalidad dentro del documento anexo para evitar la extensión desmesurada del manual central.
- **Matriz de Síntesis:** Inclusión de la tabla consolidada que cruza los ecosistemas tecnológicos, las escalas de captura asignadas y, fundamentalmente, los umbrales cuantitativos de éxito y fracaso UX. Esta matriz estandariza la interpretación de resultados empíricos para todos los proyectos.
- **Framework de Evaluación (UX-Scorecard):** Presentación de la ficha de auditoría unificada. Este instrumento constituye el entregable final del protocolo, obligando a los equipos a reportar sus hallazgos empíricos contrastados con los umbrales de éxito y derivar de ellos un plan de acción correctivo fundamentado antes del despliegue final en la FESW.

Si bien el detalle operativo y las guías de aplicación de las métricas se encuentran en el documento anexo, resulta imperativo exponer dentro del cuerpo principal de esta memoria los dos componentes estructurales que orquestan el inicio y el fin del proceso evaluativo. Por

consiguiente, a continuación se presenta in extenso la **Ruta Diagnóstica: Árbol de Decisión para la Asignación de Ecosistemas** (Sección 3.5.2), la cual actúa como el filtro de entrada al protocolo, y el **Framework de Evaluación: UX-Scorecard de Trazabilidad** (Sección 3.5.3), que constituye el entregable de salida y consolidación de resultados empíricos.

La exposición explícita de estos dos elementos permite comprender la mecánica de control y trazabilidad del protocolo propuesto, sin abrumar la lectura del capítulo con los procedimientos de cálculo estadístico detallados en el manual operativo.

3.5.2. Ruta Diagnóstica: Árbol de Decisión para la Asignación de Ecosistemas

Para garantizar la correcta adscripción de los proyectos estudiantiles a su respectiva categoría y mitigar sesgos en la selección de las métricas, el protocolo inicia con la aplicación de un Árbol de Decisión Diagnóstico. Este instrumento opera como el filtro de entrada del marco de evaluación, guiando a los equipos a través de un enrutamiento dicotómico fundamentado en las dimensiones de soporte físico y contexto operativo previamente definidas. Cabe destacar que, si bien a continuación se describe detalladamente el flujo lógico de este proceso, su estructura metodológica se encuentra esquematizada de manera íntegra en la **Figura 3.4**. Dicha representación visual en formato de árbol de decisión tiene como propósito facilitar la comprensión global y agilizar la auto-clasificación por parte de los equipos de desarrollo.

Ruta Diagnóstica:

Los equipos de desarrollo deben evaluar la naturaleza principal de la interacción de su software respondiendo de manera secuencial a los siguientes niveles lógicos:

NIVEL 1: Dimensión de Soporte Físico y Hardware Especializado

El primer nivel de filtrado determina si la interacción central del proyecto está supeditada al uso de hardware periférico, inmersión espacial o captura de datos mediante sensores.

Pregunta: ¿La interacción central de su proyecto depende del uso de hardware periférico, inmersión espacial o captura de datos mediante sensores ambientales/corporales?

■ **SÍ depende de hardware/sensores:**

El equipo debe identificar su caso de uso específico.

- **Caso A:** Utiliza cascos/visores para mundos virtuales o superpone elementos interactivos en la realidad mediante cámaras. → **Asignación: Ecosistema 5 (Inmersivo y Háptico - VR/AR).**
- **Caso B:** El dispositivo se lleva acoplado al cuerpo(ej. reloj inteligente, monitores biométricos) y requiere micro-interacciones. → **Asignación: Ecosistema 8 (Wearables y Salud Digital).**
- **Caso C:** El núcleo operativo procesa en tiempo real lo que la cámara ve o el micrófono escucha (reconocimiento facial, de gestos o voz). → **Asignación: Ecosistema 3 (Percepción Sensorial).**
- **Caso D:** Captura telemetría de redes de sensores físicos (IoT) para que un algoritmo de Inteligencia Artificial prediga anomalías o automatice respuestas, software de mantenimiento predictivo industrial, visualizadores de analítica avanzada, Automatización hogareña o urbana (Smart Cities) y plataformas de mantenimiento predictivo industrial con proyecciones y visualización de anomalías basadas en Machine Learning. → **Asignación: Ecosistema 7 (Monitoreo Predictivo e IoT).**

- **NO depende de hardware especializado:** La interacción ocurre a través de pantallas convencionales (Web, Desktop, Mobile tradicional). → *Continúe al Nivel 2.*

NIVEL 2: Dimensión de Arquitectura Lógica y Contexto Operativo *Instrucción:* Si su software utiliza interfaces gráficas tradicionales (no depende de hardware periférico), determine cuál es el factor crítico o la mayor exigencia cognitiva impuesta al usuario final.

■ **Contextos de Alta Presión o Carga Psicológica Especializada:**

- *Caso E:* El sistema audita vulnerabilidades, monitorea redes o gestiona incidentes donde un error u omisión compromete la seguridad, exigiendo respuestas bajo presión de tiempo. → **Asignación: Ecosistema 6 (Ciberseguridad y Sistemas Críticos).**

- *Caso F*: El objetivo principal del sistema es enseñar, evaluar retención de conocimientos o entrenar habilidades mediante un enganche lúdico o simulaciones académicas. → **Asignación: Ecosistema 4 (Formación Interactiva y EdTech).**

■ **Contextos de Uso Comercial, Administrativo o Productivo (Sistemas Base):**

- *Caso G*: Es una plataforma orientada a la movilidad, donde el intercambio ágil de información o el registro de operaciones debe realizarse rápidamente en pantallas reducidas (interfaz táctil y zona del pulgar) bajo contextos de atención fragmentada. → **Asignación: Ecosistema 2 (Movilidad y Servicios Transaccionales).**
- *Caso H*: Es un sistema robusto (Web/SaaS/Desktop) diseñado para procesar altos volúmenes de datos, flujos de trabajo prolongados, paneles de administración (dashboards) o investigación tecnológica profunda. → **Asignación: Ecosistema 1 (Gestión, SaaS y Big Data).**

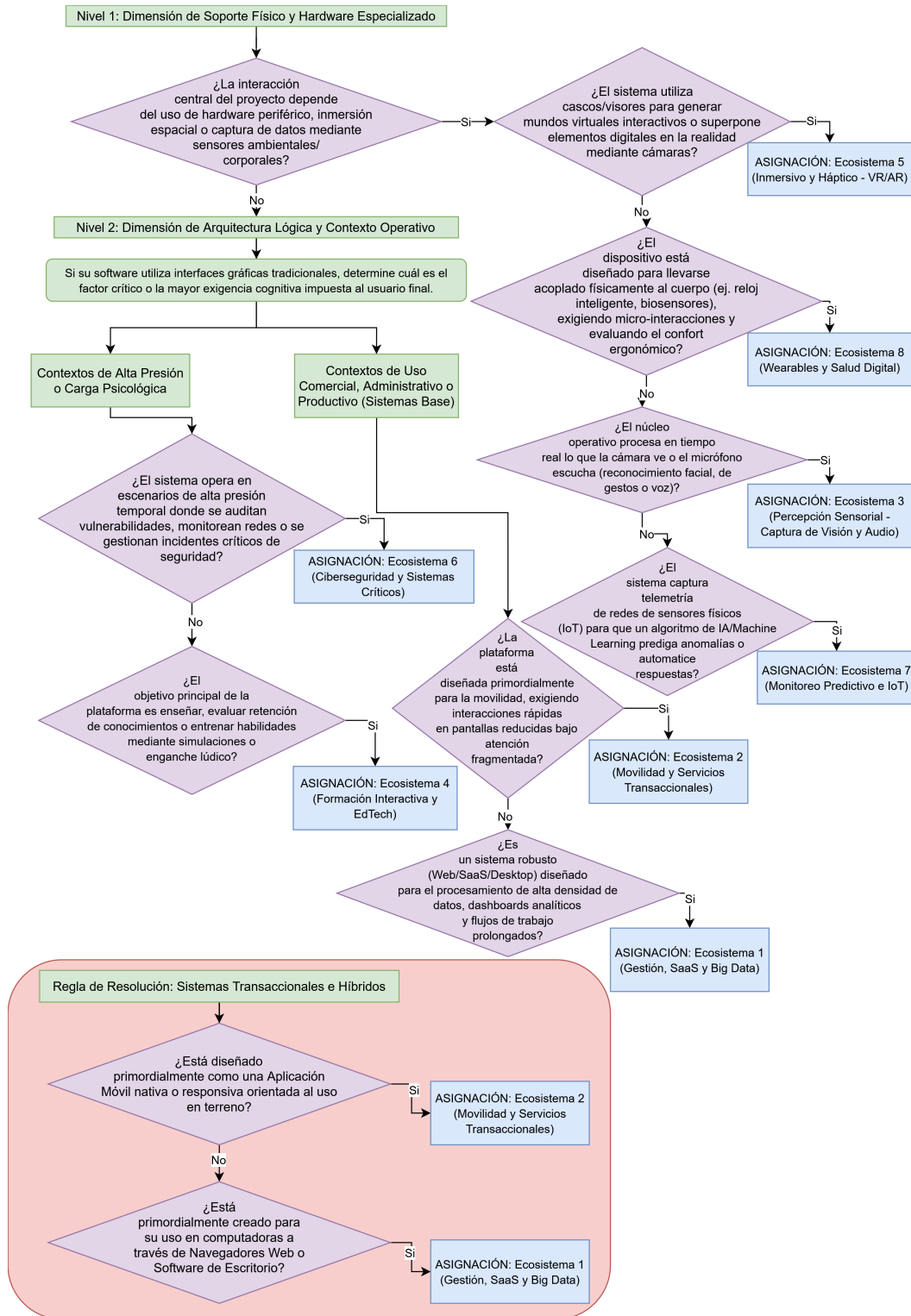
Regla de Resolución para Proyectos Híbridos (Fallback)

Si el equipo desarrollador determina que su proyecto presenta una arquitectura híbrida o sus funcionalidades no se enmarcan en las exigencias cognitivas especializadas de los Casos E y F, la solución debe ser tratada como un sistema estándar de intercambio de datos. En este escenario, aplique la siguiente regla de desempate basada en el entorno de despliegue primario:

- Si el proyecto está diseñado primordialmente como una **Aplicación Móvil** nativa o responsiva orientada al uso en terreno, clasifíquelo obligatoriamente en el **Ecosistema 2**.
- Si el proyecto está concebido primordialmente para su uso en computadoras a través de **Navegadores Web o Software de Escritorio**, clasifíquelo obligatoriamente en el **Ecosistema 1**.

La superación de esta ruta diagnóstica concluye la fase de planificación del protocolo. Una vez que el equipo de desarrollo identifica unívocamente su ecosistema, queda habilitado para adoptar la dupla de escalas psicométricas predefinida en la matriz de triangulación metodológica, dando inicio formal a la auditoría empírica de su Proyecto.

Figura 3.4: Árbol de decisión



3.5.3. Framework de Evaluación: UX-Scorecard de Trazabilidad

Una vez finalizadas las etapas de diagnóstico, ejecución empírica y tabulación matemática dictadas en el manual operativo, resulta indispensable consolidar los hallazgos. Para evitar que la evaluación de Experiencia de Usuario quede relegada a un análisis anecdótico o informal, el protocolo estandariza el reporte de resultados a través del instrumento denominado *UX-Scorecard*.

Este framework opera como una ficha de auditoría unificada y obligatoria. Su propósito metodológico es estructurar la trazabilidad de la calidad del software, obligando a los equipos a contrastar sus puntajes empíricos (micro-tácticos y macro-estratégicos) con los umbrales de éxito predefinidos (véase Tabla 3.1). A partir de este contraste, los desarrolladores deben derivar y justificar un plan de acción correctivo antes del despliegue final del proyecto en la feria.

A continuación, en el Cuadro 3.2, se presenta la estructura formal del UX-Scorecard que los equipos deben integrar como entregable oficial en la documentación de su proyecto:

Tabla 3.2: Plantilla oficial del UX-Scorecard para la evaluación de proyectos tecnológicos.

FICHA DE AUDITORÍA Y TRAZABILIDAD UX (UX-SCORECARD)	
1. IDENTIFICACIÓN DEL PROYECTO	
Nombre del Proyecto:	[Nombre del Software / Startup]
Ecosistema Asignado:	[Ej. Ecosistema 6: Ciberseguridad y Sistemas Críticos]
Hito de Evaluación:	[Ej. Cierre de Sprint 2 / Previo a Demo Day]
Tamaño de Muestra (N):	[Número de usuarios reales evaluados]
2. RESULTADOS DE TRIANGULACIÓN METODOLÓGICA	
Métrica 1 (Post-Tarea/Fricción):	[Nombre de la Escala. Ej: SEQ (Facilidad Única)]
Puntaje Obtenido:	[Puntaje numérico] / [Rango máximo de la escala]
Diagnóstico de Métrica 1:	☐ Óptimo ☐ Aceptable/Marginal ☐ Crítico (Fallo)
Métrica 2 (Post-Test/Global):	[Nombre de la Escala. Ej: Jian Trust (Confianza)]
Puntaje Obtenido:	[Puntaje numérico] / [Rango máximo de la escala]
Diagnóstico de Métrica 2:	☐ Óptimo ☐ Aceptable/Marginal ☐ Crítico (Fallo)
3. ANÁLISIS DE BRECHAS Y PLAN DE ACCIÓN	
Hallazgos Críticos:	Describe brevemente dónde ocurrió la mayor fricción cognitiva o técnica según los resultados empíricos (Ej: "La métrica SEQ arrojó 2 puntos en el flujo de configuración de alertas, evidenciando confusión en la nomenclatura").
Acciones Correctivas:	▪ Ajuste 1: [Ej. Rediseño de la jerarquía visual del panel de control] ▪ Ajuste 2: [Ej. Inclusión de un tutorial interactivo onboarding]

La implementación rigurosa de este Scorecard garantiza que la toma de decisiones sobre el diseño de interacción deje de basarse en intuiciones aisladas del equipo desarrollador. En su lugar, el proceso pasa a fundamentarse en evidencia empírica estructurada, alineando los proyectos universitarios con los estándares de calidad del software exigidos por la industria del software.

Capítulo 4

Validación

El presente capítulo detalla el proceso de validación empírica del Protocolo de Evaluación UX diseñado en la etapa anterior. Una vez estructurado el marco metodológico, el catálogo de escalas y la matriz de triangulación, resulta imperativo comprobar su viabilidad, nivel de adopción y pertinencia en un escenario real de aplicación. Para ello, la evaluación se centró en los usuarios finales del instrumento: los equipos de estudiantes de las asignaturas vinculadas a la Feria de Software de la UTFSM.

Para estructurar esta validación con rigor académico, el diseño del instrumento de recolección de datos se fundamentó en los principios del Modelo de Aceptación Tecnológica (TAM, por sus siglas en inglés). Este marco teórico permite diagnosticar cómo los usuarios perciben, aceptan y adoptan una nueva herramienta metodológica, evaluando tres dimensiones críticas:

- **Utilidad Percibida:** El grado en que los equipos consideran que la aplicación del protocolo, el uso del UX-Scorecard y sus métricas asignadas ayudarían a identificar problemas de usabilidad y mejorarían la calidad final de sus proyectos de software.
- **Facilidad de Uso Percibida:** El nivel de esfuerzo cognitivo o técnico que demanda la adopción del manual. Esto incluye evaluar si aprender a clasificar el proyecto mediante la Ruta Diagnóstica, comprender las reglas de uso y tabular los resultados en la Matriz

de Síntesis requiere un esfuerzo excesivo.

- **Intención de Uso:** La disposición activa de los estudiantes para integrar y utilizar regularmente este marco de evaluación dentro de sus *sprints* de desarrollo, en el escenario de que el protocolo se consolide como material oficial de la asignatura.

A lo largo de este capítulo se presentará el análisis descriptivo de los resultados obtenidos a partir de la encuesta aplicada a los proyectos evaluados. Posteriormente, se contrastarán los hallazgos cuantitativos de aceptación con la retroalimentación cualitativa proporcionada por los participantes. Este cruce de información permitirá identificar empíricamente los puntos de fricción en la comprensión del documento y, en consecuencia, establecer un plan de acciones correctivas para la consolidación de la versión final del protocolo.

4.1. Caracterización de la Muestra

Para llevar a cabo la validación empírica del protocolo, se seleccionó una muestra por conveniencia compuesta por los 12 equipos de desarrollo de la asignatura Taller de Desarrollo de Proyectos de Informática (INF-228) del Campus Casa Central, correspondientes al ciclo académico del presente año 2026. La aplicación del instrumento se ejecutó durante la fase de desarrollo ágil de los proyectos, permitiendo a los estudiantes contrastar el marco teórico propuesto con las necesidades reales de su software en construcción.

El perfil de la cohorte resulta de alto valor metodológico, dado que abarca una amplia diversidad de arquitecturas lógicas y de hardware. Al someter sus proyectos a la Ruta Diagnóstica (Árbol de Decisión) estipulada en el manual, los 12 equipos participantes lograron autoclasificarse en siete de los ecosistemas tecnológicos definidos en la presente memoria:

Figura 4.1: Auto-clasificación de proyectos

¿A qué ecosistema tecnológico pertenece tu proyecto actual según la Ruta Diagnóstica (Árbol de Decisión) del protocolo?

12 respuestas



Nota: El Ecosistema 5 (Entornos Inmersivos y Hápticos) no contó con representantes en la presente cohorte del Campus Casa Central, lo cual refleja las fluctuaciones naturales en la adopción de hardware año a año.

Adicionalmente, el instrumento evaluó el nivel de madurez metodológica de los equipos como línea base del estudio. Al ser consultados sobre si poseían claridad respecto a qué métricas exactas utilizar para evaluar la Experiencia de Usuario antes de revisar el protocolo, el 100 % de la muestra indicó carecer de un conocimiento consolidado: 8 equipos (66,7 %) declararon no tener claridad absoluta, mientras que los 4 restantes (33,3 %) indicaron poseer solo una claridad parcial. Este antecedente empírico confirma la existencia de la brecha formativa descrita en el planteamiento del problema y fundamenta la pertinencia de la intervención metodológica en la asignatura.

4.2. Resultados Cuantitativos: Modelo de Aceptación Tecnológica

A partir de la caracterización inicial, se procedió a evaluar la adopción del instrumento mediante las 7 afirmaciones estructuradas bajo el modelo TAM, medidas a través de una escala Likert de 5 puntos (donde 1 representa “Totalmente en desacuerdo” y 5 “Totalmente de acuerdo”).

Tabla 4.1: Estadística descriptiva de la evaluación del protocolo basada en el modelo TAM (N=12).

Ítem Evaluado	1	2	3	4	5	Promedio
Dimensión: Utilidad Percibida						
1. Identificar problemas ignorados	1	0	1	2	8	4.33
2. Mejorar calidad del producto	0	0	2	1	9	4.58
3. Justificar UX-Scorecard objetivamente	0	0	1	3	8	4.58
Dimensión: Facilidad de Uso Percibida						
4. Claridad en el Árbol de Decisión	1	0	2	3	6	4.08
5. Comprensión de reglas y métricas	0	0	1	7	4	4.25
6. Esfuerzo técnico en tabular y Matriz	0	3	2	3	4	3.67
Dimensión: Intención de Uso						
7. Uso activo en <i>sprints</i> del equipo	0	3	0	4	5	3.92

Nota: Escala Likert donde 1 = Totalmente en desacuerdo y 5 = Totalmente de acuerdo.

4.2.1. Análisis Descriptivo de la Aceptación Tecnológica

El procesamiento de los datos recolectados en la Tabla 4.1 permite estructurar un diagnóstico empírico sobre la adopción del protocolo, desagregado en las tres dimensiones del modelo TAM:

1. Utilidad Percibida (Alta Aceptación): Esta dimensión obtuvo los resultados más favorables del instrumento, promediando un 4.50 sobre 5.00. Destacan particularmente los ítems referidos al impacto directo en el producto (Ítem 2) y la capacidad del UX-Scorecard para justificar objetivamente el rediseño (Ítem 3), ambos con un promedio de 4.58 y concentrando el 75 % de sus respuestas en la calificación máxima (“Totalmente de acuerdo”). Estos indicadores validan empíricamente el propósito central del protocolo generado en la presente memoria: los estudiantes reconocen que la aplicación de métricas especializadas según su ecosistema eleva la calidad del software.

2. Facilidad de Uso Percibida (Fricción Cognitiva Moderada): La percepción sobre el esfuerzo requerido para aplicar el manual promedió un 4.00 global, evidenciando mayores fricciones de interacción. Si bien la comprensión de las reglas (Ítem 5, promedio 4.25) y la

clarificación del Árbol de Decisión (Ítem 4, promedio 4.08) mantienen índices de aceptación positivos, el Ítem 6 representó el punto más crítico de toda la evaluación. Al consultar si la tabulación de resultados y el uso de la “Matriz de Síntesis” no requerían un esfuerzo mental excesivo, el promedio descendió a 3.67, registrándose un 25 % de respuestas (3 usuarios) en evidente desacuerdo (calificación 2). Este hallazgo cuantitativo corrobora que la traducción de datos crudos hacia umbrales estadísticos representa una barrera cognitiva significativa para los estudiantes de pregrado.

3. Intención de Uso (Adopción Polarizada): La proyección sobre el uso continuo del protocolo en los *sprints* de desarrollo (Ítem 7) obtuvo el segundo promedio más bajo del estudio (3.92). Si bien 9 de los 12 equipos (75 %) se mostraron “De acuerdo” o “Totalmente de acuerdo” con adoptarlo activamente si fuera material oficial, 3 equipos manifestaron rechazo (calificación 2). Esta polarización sugiere que, aunque el protocolo es percibido como una herramienta de alta utilidad técnica, factores externos al manual tales como la carga académica existente, la extensión del documento o su introducción tardía en el flujo de la asignatura generan resistencia hacia su adopción espontánea.

4.2.2. Análisis Cualitativo y Oportunidades de Mejora

La recolección de datos cualitativos mediante preguntas abiertas permitió profundizar en las percepciones de los estudiantes respecto a la implementación del protocolo. Si bien la utilidad técnica del instrumento fue ampliamente reconocida, los comentarios revelaron fricciones operativas que explican los puntajes más bajos obtenidos en las dimensiones de Facilidad de Uso e Intención de Uso del modelo TAM.

Los hallazgos cualitativos derivados de la retroalimentación de los participantes se agruparon en cinco ejes críticos:

- **Arquitectura de la Información y Navegabilidad:** Diversos estudiantes manifestaron sentirse abrumados por la extensión del documento y la dificultad para navegar hacia las rúbricas de evaluación. Se evidenció resistencia a la lectura de manuales extensos (mencionando explícitamente “no hacerles leer un protocolo de 33 páginas”) y

se reportaron confusiones respecto a “las direcciones de los archivos dentro del protocolo” o “cómo llegar a los anexos”.

- **Terminología y Carga Cognitiva:** Se reportó que el lenguaje utilizado resulta “muy técnico en las partes donde el usuario final evalúa la interfaz”. Conceptos estadísticos o teóricos como “carga nula” o la interpretación de la escala Likert generaron confusión. Los participantes sugirieron reiteradamente el uso de un “lenguaje más natural” y menos tecnicismos para facilitar la comprensión.
- **Clasificación de Ecosistemas Híbridos:** El Árbol de Decisión presentó limitaciones ante proyectos con arquitecturas superpuestas. Participantes destacaron la dificultad de enmarcar desarrollos que mezclan elementos educativos (EdTech) con plataformas móviles de uso transaccional, señalando que “los cuatro casos no son mutuamente excluyentes” y que su proyecto encajaba en más de un ecosistema simultáneamente.
- **Adaptabilidad Heurística:** Se observó que las heurísticas de usabilidad universales (como los principios de Nielsen) requieren contextualización según el soporte físico. Un estudiante evidenció este desajuste al mencionar que la heurística referente a “atajos” o “comandos de teclado” no calza adecuadamente con el paradigma de interacción de los dispositivos móviles.
- **Integración Temprana y Automatización:** Los equipos manifestaron que el protocolo se introdujo tarde en su ciclo académico (“nos llegó más bien tarde”), sugiriendo que estas directrices debiesen implementarse en principio del primer semestre. Adicionalmente, sugirieron que la asimilación del protocolo podría facilitarse integrándolo al flujo de trabajo (*workflow*) mediante herramientas de Inteligencia Artificial que sistematicen la selección de métricas o generen tareas en plataformas de gestión.

Para consolidar la versión definitiva del manual operativo y asegurar su viabilidad en la asignatura, se elaboró una matriz de acciones correctivas. La Tabla 4.2 vincula directamente los hallazgos empíricos de la validación con los ajustes metodológicos aplicados al protocolo y las proyecciones de trabajo futuro.

Tabla 4.2: Matriz de Hallazgos Cualitativos y Plan de Acción Correctivo del Protocolo UX.

Dimensión Evaluada	Evidencia Cualitativa (Hallazgo Principal)	Acción Correctiva Implementada en el Protocolo y Memoria
Arquitectura de la Información	Dificultad para navegar el documento y sobrecarga cognitiva por su extensión (33 páginas).	Reestructuración Modular: Separación estricta del marco teórico de las guías de uso operativo. Incorporación de hipervínculos internos interactivos en el manual final para saltar directamente desde el Árbol de Decisión al anexo correspondiente.
Clasificación (Árbol de Decisión)	Los niveles de diagnóstico resultan excluyentes para proyectos con arquitecturas híbridas (ej. Móvil + EdTech).	Regla de Resolución (3.5.2 Ruta Diagnóstica - Fallback): Incorporación de una instrucción explícita que prioriza la clasificación de sistemas híbridos hacia su base transaccional de despliegue primario (Escritorio o Móvil).
Lenguaje y Accesibilidad	Exceso de tecnicismos (“escala Likert”, “carga nula”) que dificultan la comprensión del usuario final.	Simplificación y Glosario: Redacción de instrucciones en lenguaje natural para los usuarios de prueba, y consolidación de una nota aclaratoria temprana que define conceptos de software.
Adaptabilidad Heurística (Nielsen)	Los principios universales presentan desajustes al auditar soportes físicos específicos como aplicaciones móviles.	Nota de Adaptación Contextual: Adición de directrices para adaptar las heurísticas tradicionales (ej. reemplazar la evaluación de atajos de teclado por la de gestos táctiles en entornos móviles).
Integración al Flujo de Trabajo	La adopción del manual requiere integración temprana (primer semestre) y herramientas de gestión ágil o IA.	Proyección Tecnológica: Proponer en el apartado de “Trabajo Futuro” la automatización de este protocolo mediante agentes de Inteligencia Artificial para integrarlo directamente a los <i>sprints</i> de los equipos.

4.2.3. Validación Final: Prueba de Concepto (Estudios de Caso)

Posterior a la estructuración de la matriz de acciones correctivas, se ejecutó una segunda fase de validación empírica. El objetivo de esta etapa fue comprobar la eficacia técnica y la facilidad de adopción de las modificaciones implementadas en el manual operativo, aplicando

todos los ajustes propuestos con la única excepción del ítem “Integración al Flujo de Trabajo”, el cual, por su complejidad técnica de automatización, fue relegado formalmente como proyección para el Trabajo Futuro de esta investigación.

Para llevar a cabo esta Prueba de Concepto mediante Estudios de Caso, se seleccionaron intencionalmente dos de los proyectos de la cohorte que experimentaron la mayor fricción cognitiva y declararon las mayores confusiones durante la evaluación inicial. El propósito de seleccionar los escenarios más críticos fue someter la nueva versión del documento a una prueba de estrés metodológico. Los proyectos seleccionados fueron VetaSearch y Zenithy. El primero corresponde a un sistema orientado al análisis de información del entorno minero, con el propósito de identificar y anticipar potenciales situaciones de riesgo. El segundo, Zenithy, consiste en una aplicación enfocada en el monitoreo del bienestar de los usuarios mediante el uso de inteligencia artificial y herramientas de engagement. A estos dos equipos se les presentó la versión depurada y actualizada del protocolo UX. Se les instruyó seguir el flujo metodológico completo: transitar de manera autónoma por la Ruta Diagnóstica para auto-clasificar su sistema, adoptar las escalas psicométricas exactas asignadas a su ecosistema, y finalmente, aplicar estos instrumentos empíricos a un usuario real de prueba durante una sesión de *testing*.

Finalizada la sesión práctica, se realizó una entrevista de cierre con los integrantes de los proyectos para evaluar la utilidad del instrumento actualizado utilizando la misma encuesta realizada anteriormente. Además se les consultó directamente si la aplicación de las métricas les permitió hallar fricciones o problemas de interacción que requiriesen modificaciones en sus interfaces y ambos grupos respondieron afirmativamente, validando la utilidad diagnóstica de la herramienta.

Tabla 4.3: Estadística descriptiva de la evaluación del protocolo basada en el modelo TAM (N=2).

Ítem Evaluado	1	2	3	4	5	Promedio
Dimensión: Utilidad Percibida						
1. Identificar problemas ignorados	0	0	0	0	2	5.0
2. Mejorar calidad del producto	0	0	0	0	2	5.0
3. Justificar UX-Scorecard objetivamente	0	0	0	0	2	5.0
Dimensión: Facilidad de Uso Percibida						
4. Claridad en el Árbol de Decisión	0	0	0	1	1	4.50
5. Comprensión de reglas y métricas	0	0	0	0	2	5.0
6. Esfuerzo técnico en tabular y Matriz	0	0	0	1	1	4.50
Dimensión: Intención de Uso						
7. Uso activo en <i>sprints</i> del equipo	0	0	0	0	2	5

Nota: Escala Likert donde 1 = Totalmente en desacuerdo y 5 = Totalmente de acuerdo.

Al contrastar estos datos empíricos con la medición inicial (véase Tabla 4.1), se evidencia una mejora sustancial en todas las dimensiones. La **Facilidad de Uso Percibida**, que representaba la principal barrera cognitiva (ítem 6 con promedio inicial de 3.67), se elevó a 4.50, confirmando que la incorporación de un glosario y lenguaje natural erradicaron la confusión técnica. Asimismo, la **Intención de Uso** alcanzó la puntuación máxima perfecta (5.00), demostrando que al optimizar la arquitectura de la información y clarificar los proyectos híbridos, los equipos están dispuestos a adoptar la herramienta sin resistencia.

Una vez validada la facilidad de uso del documento, se procedió a realizar una entrevista de cierre con los integrantes de ambos proyectos para evaluar la capacidad diagnóstica real del instrumento. Se les consultó directamente si, tras tabular los resultados obtenidos de sus usuarios reales, la aplicación de las métricas les permitió hallar fricciones o problemas de interacción ocultos que requiriesen modificaciones en sus interfaces. Ambos grupos respondieron afirmativamente.

Específicamente, el equipo desarrollador del proyecto *VetaSearch* destacó que el uso de las escalas de evaluación (pertenecientes al Ecosistema 1: Gestión, SaaS y Big Data (Web/Desktop

) [3.4.2.1]) les permitió identificar un sesgo cognitivo y un cuello de botella en su arquitectura de información. Tras aplicar el instrumento, el equipo notó que la disposición visual de su interfaz hacía sentir al usuario que debía abrir un documento que, en realidad, ya se encontraba abierto, generándole confusión y pérdida de tiempo. Los desarrolladores concluyeron que la aplicación del protocolo les ayudó a encontrar detalles críticos en la interfaz que, al ser corregidos, asegurarán que el usuario final interactúe con mucha mayor fluidez.

Por otro lado, el equipo del proyecto *Zenithy* perteneciente al Ecosistema 2: Movilidad y Servicios Transaccionales [3.4.2.1] reportó que la escala SUPR-Qm resultó fundamental para entender y diagnosticar problemas latentes en su aplicación. En particular, destacaron la relevancia de la dimensión de Confianza, evaluada a través de reactivos como “*Siento que mis datos e información personal están seguros en esta plataforma*”. Dado que el objetivo primordial de su proyecto exige que el usuario pueda confiar plenamente en el manejo de sus datos, la asignación estratégica de esta métrica les permitió auditar y robustecer la seguridad percibida en su diseño, algo que una evaluación de usabilidad genérica no habría logrado capturar.

La culminación exitosa de esta segunda iteración demuestra empíricamente que el protocolo funciona, es adaptable a escenarios complejos y otorga resultados reales y tangibles para el aseguramiento de la calidad de la Experiencia de Usuario.

Para documentar la ejecución empírica de esta fase, las Figuras 4.2 y 4.3 ilustran el entorno de *testing* moderado (en el laboratorio de Experiencia Usuaría de UTFSM Casa Central). En estas instancias, los usuarios finales interactuaron con las plataformas en tiempo real, permitiendo a los desarrolladores aplicar las métricas post-tarea y post-test dictadas por el manual, y contrastar los puntajes obtenidos con el comportamiento gestual y la fricción observada durante la prueba.

Figura 4.2: Proyecto VetaSearch

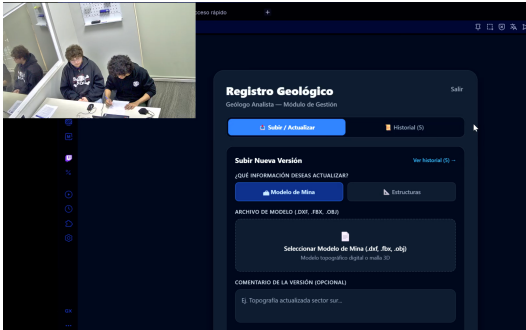
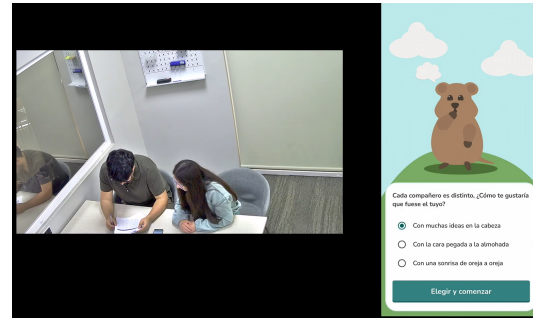


Figura 4.3: Proyecto Zenithy



4.2.4. Síntesis y Cierre de la Validación Empírica

En conclusión, el proceso de validación iterativa documentado en este capítulo comprueba de manera exhaustiva la pertinencia, usabilidad y eficacia del Protocolo de Evaluación UX diseñado. La primera fase de diagnóstico, sustentada en el Modelo de Aceptación Tecnológica (TAM), permitió identificar y erradicar oportunamente las barreras cognitivas y de arquitectura de la información que dificultaban la adopción del manual por parte de los estudiantes.

Posteriormente, la Prueba de Concepto mediante los estudios de caso demostró que, al operar con la versión depurada del instrumento, los equipos logran ejecutar auditorías de usabilidad con rigor metodológico. El protocolo guía exitosamente a los estudiantes en la aislación de fricciones críticas tanto de interacción visual como de confianza algorítmica que las pruebas de software puramente funcionales suelen omitir.

En consecuencia, la herramienta metodológica propuesta en esta memoria trasciende la categoría de marco teórico para consolidarse empíricamente como un instrumento operativo viable, flexible y directamente integrable en el ciclo de desarrollo ágil de la Feria de Software.

Conclusiones y trabajo futuro

El presente trabajo de título se desarrolló con el propósito de resolver una brecha metodológica crítica en la formulación y gestión de proyectos informáticos dentro de las asignaturas que culminan en la Feria de Software de la Universidad Técnica Federico Santa María. Históricamente, la postergación de los atributos de usabilidad frente a la urgencia de la implementación técnica resultaba en soluciones que, si bien operaban correctamente a nivel de código, presentaban fricciones de interacción que limitaban su adopción y el cumplimiento de sus Objetivos de Desarrollo Sostenible (ODS).

A continuación, se presentan las conclusiones derivadas de la investigación, estructuradas inicialmente en torno a la evaluación del grado de cumplimiento de los objetivos propuestos, seguidas por una reflexión sobre el impacto metodológico del protocolo y las proyecciones para trabajos futuros.

5.1. Conclusiones Generales

5.1.1. Cumplimiento de los Objetivos de la Investigación

El **Objetivo General** de esta memoria consistió en diseñar un protocolo de evaluación de Experiencia de Usuario mediante métricas y escalas estandarizadas, adaptado a los múltiples ecosistemas tecnológicos presentes en la Feria de Software. Tras el desarrollo y la validación de la propuesta, se concluye que este objetivo fue alcanzado a cabalidad, lo cual se fundamenta en el cumplimiento secuencial de cada uno de los objetivos específicos trazados al

inicio del proyecto:

- **Investigación de métricas y escalas estandarizadas:** Se completó exitosamente a través de una exhaustiva revisión del estado del arte (Capítulo 2). Se logró identificar y analizar un catálogo robusto de instrumentos psicométricos y de auditoría experta (tales como SUS, SUPR-Qm, GUESS-18, PQ, SSQ, CRS, Jian Trust, NASA-TLX, UMUX-Lite PURE y SEQ), evaluando rigurosamente su pertinencia científica frente a diversos entornos operativos.
- **Categorización preliminar de proyectos:** Se logró establecer una taxonomía tecnológica inicial (Paso 1, Capítulo 3) fundamentada en la dimensión del soporte físico y la arquitectura lógica, analizando una muestra empírica de 30 proyectos presentados en la cohorte 2024 y sumando el conocimiento observacional de tendencias emergentes en la industria.
- **Validación y refinamiento de la taxonomía:** Se cumplió íntegramente mediante el levantamiento de información empírica con 72 participantes de ediciones anteriores del evento (2019-2025). Esta validación de pares permitió robustecer el modelo, integrando vacíos metodológicos detectados empíricamente, como la necesidad de independizar los sistemas de ciberseguridad, el procesamiento de audio y diferenciar el software EdTech de los videojuegos tradicionales.
- **Asignación estratégica mediante triangulación metodológica:** Se materializó exitosamente a través de la creación de la “Matriz de Síntesis” (Paso 4, Capítulo 3.4). La estructuración de duplas de evaluación permitió cruzar métricas micro-tácticas (fricción post-tarea) con métricas macro-estratégicas (percepción post-test), aislando factores críticos de frontera como la fatiga ergonómica o la confianza algorítmica según la naturaleza del ecosistema.
- **Diseño del protocolo estructurado:** Se cumplió mediante la construcción del “Protocolo de Evaluación UX para proyectos desarrollados para la Feria de Software”, el cual articula el Árbol de Decisión Diagnóstico, las fichas operativas por ecosistema, y estandariza la toma de decisiones a través del *Framework* de Trazabilidad denominado *UX-Scorecard*.

- **Validación final del protocolo:** El objetivo se completó satisfactoriamente en el Capítulo 4. La aplicación del Modelo de Aceptación Tecnológica (TAM) a 12 equipos de la cohorte 2026, seguida de una Prueba de Concepto iterativa con los proyectos (*Veta-search* y *Zenithy*), permitió depurar las fricciones del documento original y confirmar empíricamente su utilidad diagnóstica, su facilidad de adopción y la alta intención de uso por parte del alumnado.

En síntesis, la articulación de estas etapas demuestra que la carencia de métricas adaptadas para la evaluación heterogénea de proyectos ha sido cubierta, dotando a las asignaturas de un marco de trabajo sólido, replicable y fundamentado en estándares internacionales de Interacción Humano-Computadora.

5.1.2. Reflexiones Metodológicas y Prácticas

La implementación de este protocolo de evaluación establece un cambio de paradigma en el control de calidad de los proyectos desarrollados en las asignaturas formativas. En primer lugar, la adopción de un modelo de triangulación metodológica demostró ser el pilar fundamental para abordar la alta heterogeneidad tecnológica de las propuestas. Históricamente, la aplicación de encuestas de usabilidad genéricas a sistemas dispares generaba diagnósticos sesgados, equiparando escenarios operativos incomparables. Al cruzar métricas micro-tácticas (fricción inmediata) con evaluaciones macro-estratégicas adaptadas a la plataforma, el protocolo evita el error metodológico de medir con la misma regla exigencias cognitivas distintas. Esta focalización permite aislar variables críticas de frontera como la fatiga biomecánica en un dispositivo vestible o la desconfianza algorítmica en un panel de ciberseguridad garantizando que las métricas respondan a la naturaleza técnica y física real de cada ecosistema.

En segundo lugar, la estandarización del proceso analítico a través del *Framework UX-Scorecard 3.2* aporta un inmenso valor práctico al ciclo de desarrollo ágil de los equipos. Tradicionalmente, las decisiones de rediseño de interfaces en instancias académicas tienden a fundamentarse en la intuición de los desarrolladores o en la mera corrección de *bugs* de

código, postergando la experiencia del usuario final. La exigencia del protocolo de tabular resultados y contrastarlos con los umbrales de éxito predefinidos en la Matriz de Síntesis obliga a los estudiantes a justificar sus iteraciones basándose estrictamente en evidencia empírica.

De este modo, el *UX-Scorecard* 3.2 opera como un mecanismo de objetividad: traduce la frustración subjetiva del usuario en datos cuantitativos, procesables y accionables. Esto asegura que el esfuerzo de programación invertido en los *sprints* finales responda a la solución de cuellos de botella validados por el mercado objetivo, elevando la madurez de los proyectos universitarios hacia los estándares de calidad exigidos por la industria del software.

5.2. Trabajo Futuro

Como proyección natural de esta investigación, el trabajo futuro se orienta a la sistematización y automatización del protocolo diseñado. Si bien el manual operativo y el *UX-Scorecard* 3.2 han demostrado su viabilidad empírica y utilidad diagnóstica, la ejecución de la Ruta Diagnóstica y la tabulación de resultados aún demandan un esfuerzo manual por parte de los equipos de desarrollo. Para reducir al mínimo esta carga operativa y asegurar una adopción nativa dentro de las asignaturas, la siguiente etapa metodológica debiese consistir en la digitalización interactiva del instrumento.

Específicamente, se propone el desarrollo de un agente de Inteligencia Artificial (IA) pre-entrenado bajo la taxonomía, reglas de hibridación y umbrales de éxito definidos en esta memoria. El propósito central de esa implementación tecnológica es facilitar drásticamente la experiencia de los estudiantes: al ingresar la descripción de su software, los equipos obtendrían de manera inmediata y automatizada el ecosistema al que pertenece su proyecto y las escalas exactas que deben utilizar, eliminando por completo la necesidad de transitar manualmente por el Árbol de Decisión.

Además de facilitar el diagnóstico metodológico, este asistente virtual podría integrarse de manera directa en los flujos de trabajo (*workflows*) de los alumnos. Mediante conexiones API, la herramienta generaría y asignaría automáticamente los *tickets* de evaluación empírica

en plataformas de gestión ágil como Jira, Trello o Notion. La materialización de esta integración transformaría el presente protocolo estático en un componente dinámico del *pipeline* de desarrollo, consolidando definitivamente la evaluación de la Experiencia de Usuario como un proceso ágil, ineludible y altamente eficiente en la Feria de Software.

Bibliografía

- [1] International Organization for Standardization, “Ergonomics of human-system interaction — part 210: Human-centred design for interactive systems,” 2019.
- [2] D. Norman and J. Nielsen, “The definition of user experience (ux),” 8 1998.
- [3] B. Sheppard, G. Kouyoumjian, H. Sarrazin, and F. Dore, “The business value of design,” *McKinsey Quarterly*, oct 2018.
- [4] R. S. Pressman and B. R. Maxim, *Software Engineering: A Practitioner’s Approach*. New York, NY: McGraw-Hill Education, 9th ed., 2020.
- [5] A. J. Dutson, R. H. Todd, S. P. Magleby, and C. D. Sorensen, “A review of literature on teaching engineering design through project-oriented capstone courses,” *Journal of Engineering Education*, vol. 86, no. 1, pp. 17–28, 1997.
- [6] J. Nielsen, “Usability engineering,” in *Usability Engineering*, pp. 1–362, Academic Press/AP Professional, 1993.
- [7] B. Shneiderman, C. Plaisant, M. Cohen, S. Jacobs, N. Elmqvist, and N. Diakopoulos, *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Boston: Pearson, 6 ed., 2017.
- [8] A. Martín, G. Gaetán, V. Saldaño, C. Cardozo, A. Carrizo, H. Sosa, R. Piñero, and S. Villagra, “La experiencia de usuario en el contexto de los multi-dispositivos,” *Informes Científicos Técnicos - UNPA*, vol. 11, no. 3, pp. 111–134, 2019.
- [9] M. K. López Asto, J. J. Vásquez Cerna, and A. C. Mendoza de los Santos, “El impacto del design thinking en la tecnología,” *Revista Innovación y Software*, vol. 4, no. 2, pp. 63–78, 2023.
- [10] J. Filippi, R. Bertone, and R. García-Martínez, “Certificación de procesos y estandarización flexible de la calidad,” *Investigación Joven*, vol. 6, no. Especial, pp. 168–169, 2019.

- [11] Y. Barroso Benítez, Y. Trujillo Casañola, and Y. Millet Lombida, “Marco de trabajo de evaluación de experiencia de usuario en el desarrollo de software,” *Revista Cubana de Ciencias Informáticas*, vol. 15, no. 3, pp. 92–117, 2021.
- [12] J. Sauro and J. R. Lewis, *Quantifying the User Experience: Practical Statistics for User Research*. Waltham, MA: Morgan Kaufmann, 2012. Includes discussion on the Standardized User Experience Percentile Rank Questionnaire.
- [13] N. K. Denzin, *The Research Act: A Theoretical Introduction to Sociological Methods*. New York, NY, USA: McGraw-Hill, 2nd ed., 1978.
- [14] J. Brooke, “Sus - a quick and dirty usability scale,” in *Usability Evaluation In Industry*, pp. 189–194, Taylor & Francis, 1996.
- [15] K. Rodden, H. Hutchinson, and X. Fu, “Measuring the user experience on a large scale: user-centered metrics for web applications,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2395–2398, ACM, 2010.
- [16] J. Sauro, “Supr-m: A standardized questionnaire for measuring the mobile user experience,” *MeasuringU*, 2016. Publicado en línea en MeasuringU.com.
- [17] M. H. Phan, *The development and validation of the Game User Experience Satisfaction Scale (GUESS)*. Dissertation, Wichita State University, Wichita, KS, 2015.
- [18] B. G. Witmer and M. J. Singer, “Measuring presence in virtual environments: A presence questionnaire,” *Presence: Teleoperators & Virtual Environments*, vol. 7, no. 3, pp. 225–240, 1998.
- [19] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, “Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness,” *The International Journal of Aviation Psychology*, vol. 3, no. 3, pp. 203–220, 1993.
- [20] J. F. Knight and C. Baber, “A tool to assess the comfort of wearable computers,” *Human Factors*, vol. 47, no. 1, pp. 77–91, 2005.
- [21] E. M. Macías Vélez and E. A. Llor Navia, “Propuesta de diseño de una metodología de desarrollo de software basado en el enfoque ui/ux,” *Código Científico Revista de Investigación*, vol. 4, no. E2, pp. 387–405, 2023.
- [22] J.-Y. Jian, A. M. Bisantz, C. G. Drury, and J. Llinas, “Foundations for an empirically determined scale of trust in automated systems,” in *AFRL-HE-WP-TR-2000-0102 Interim Report*, pp. 1–35, United States Air Force Research Laboratory, 1998.
- [23] J. Sauro and J. S. Dumas, “Comparison of three one-question, post-task usability questionnaires,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*, (Boston, MA, USA), pp. 1599–1608, ACM, 2009.

- [24] S. G. Hart and L. E. Staveland, “Development of nasa-tlx (task load index): Results of empirical and theoretical research,” in *Human Mental Workload*, pp. 139–183, North-Holland: Elsevier Science Publishers, 1988.
- [25] J. Sauro, “Pure: A method to measure usability magnitude,” *MeasuringU*, 2016. Published online at MeasuringU.com.
- [26] J. R. Lewis, B. S. Utesch, and D. E. Maher, “Umux-lite – when there’s no time for the sus,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’13)*, (Paris, France), ACM, 2013.

Apéndice A

Protocolo y Guías de Uso de las Escalas de Evaluación UX

El presente anexo contiene los enlaces directos al repositorio digital en Google Drive donde se aloja el “*Protocolo de Evaluación UX para proyectos desarrollados para la Feria de software*” y las guías de uso detalladas, instrucciones de tabulación y formatos de los instrumentos psicométricos utilizados en el protocolo.

Se puede acceder al protocolo haciendo clic en el anexo siguiente:

Anexo A: Protocolo de Evaluación UX para proyectos desarrollados para la Feria de software.

Se puede acceder a la documentación completa de las guías de uso de cada escala de evaluación haciendo clic en el nombre de la escala correspondiente:

Anexo B: Guía de Uso Escala CRS con Anova. Instrumento diseñado para evaluar el confort ergonómico y el impacto psicológico en dispositivos vestibles dinámicos.

Anexo C: Guía de Uso Escala GUESS-18. Herramienta validada psicométricamente para evaluar nueve dimensiones clave de satisfacción en entornos de entretenimiento y formación interactiva.

Anexo D: Guía de Uso Escala Jian Trust. Cuestionario orientado a diagnosticar los niveles de confianza y desconfianza del usuario frente a sistemas automatizados y algoritmos.

Anexo E: Guía de Uso Escala NASA-TLX. Instrumento para diagnosticar el costo cognitivo, físico y emocional al ejecutar tareas bajo alta presión en sistemas críticos.

Anexo F: Guía de Uso Escala PQ. Escala diseñada para cuantificar la inmersión psicológica y la sensación de “estar allí” en entornos de Realidad Virtual y Aumentada.

Anexo G: Guía de Uso Método PURE. Técnica de inspección analítica para predecir y cuantificar la fricción cognitiva en interfaces sin requerir usuarios finales.

Anexo H: Guía de Uso Escala SEQ. Métrica micro-táctica post-tarea diseñada para medir la fricción inmediata generada durante la ejecución de una actividad específica.

Anexo I: Guía de Uso Escala SSQ. Instrumento diagnóstico enfocado en detectar fricciones y desgaste fisiológico generado por el hardware inmersivo, evaluando factores oculomotores, desorientación y náuseas.

Anexo J: Guía de Uso Escala SUPR-Qm. Diagnóstico holístico para interfaces móviles que evalúa usabilidad, lealtad, credibilidad y atractivo visual.

Anexo K: Guía de Uso Escala SUS. Herramienta global estandarizada para medir de manera rápida la usabilidad percibida en plataformas transaccionales complejas.

Anexo L: Guía de Uso Escala UMUX-LITE. Instrumento psicométrico ultra-breve que evalúa la percepción global de la utilidad y la facilidad de uso del sistema.