

# Informe Final - Mapeo Sistemático sobre Calidad de Software

Kevin Carrasco  
William Chimborazo  
Luis Quishpe  
Ismael Sailema

9 de septiembre de 2025

## Índice

|                                                     |           |
|-----------------------------------------------------|-----------|
| <b>1. Resumen</b>                                   | <b>2</b>  |
| <b>2. Palabras Clave</b>                            | <b>2</b>  |
| <b>3. Introducción</b>                              | <b>2</b>  |
| <b>4. Metodología</b>                               | <b>3</b>  |
| 4.1. Protocolo . . . . .                            | 3         |
| 4.2. Bases de datos y cadenas de búsqueda . . . . . | 5         |
| 4.3. Criterios de inclusión y exclusión . . . . .   | 5         |
| <b>5. Resultados del Mapeo</b>                      | <b>6</b>  |
| 5.1. Datos recopilados . . . . .                    | 6         |
| 5.2. Visualizaciones . . . . .                      | 8         |
| 5.3. Taxonomía . . . . .                            | 11        |
| <b>6. Discusión</b>                                 | <b>13</b> |
| <b>7. Amenazas a la Validez</b>                     | <b>15</b> |
| <b>8. Conclusiones y Trabajo Futuro</b>             | <b>16</b> |
| <b>Referencias</b>                                  | <b>17</b> |

# 1. Resumen

Este mapeo sistemático identifica, clasifica y visualiza el panorama de investigación sobre calidad de software con foco en: (1) herramientas de testing, (2) generación y priorización de casos de prueba y (3) métricas y evidencia empírica, alineando los aportes con las ocho características del modelo ISO/IEC 25010. El alcance es conceptual y se centra en cómo las técnicas y herramientas se vinculan con las dimensiones de calidad y cómo se operacionalizan mediante métricas. La metodología se basó en un protocolo con enfoque PICOC, búsqueda en bases especializadas (IEEE Xplore, ACM DL, Scopus) y repositorios en español (SciELO, Redalyc, Dialnet), cribado por título/resumen y texto completo, extracción estandarizada y síntesis descriptiva.

Principales hallazgos (a partir del piloto inicial): (a) predominan los aportes orientados a mantenibilidad, usabilidad y eficiencia de desempeño; (b) compatibilidad y portabilidad muestran menor cobertura; (c) se observan métricas recurrentes como cobertura, "mutation score", complejidad/acoplamiento, latencias p95/p99 y cuestionarios de usabilidad (SUS/UMUX-Lite); (d) la evidencia es mayoritariamente académica, con disponibilidad limitada de artefactos reproducibles e integración en CI/CD. El resultado incluye una taxonomía de aportes vs. características ISO/IEC 25010, visualizaciones de síntesis y un conjunto de referencias en acceso abierto para apoyar nuevas investigaciones y la práctica.

# 2. Palabras Clave

Calidad de software; ISO/IEC 25010; Testing de software; Generación y priorización de pruebas; Métricas; Evidencia empírica.

# 3. Introducción

La calidad del software es un atributo transversal que condiciona la confiabilidad, el rendimiento y la experiencia de uso de los sistemas, y que impacta directamente en los costos de mantenimiento, el riesgo operativo y la adopción por parte de los usuarios. Para enmarcar de manera rigurosa este concepto, el modelo ISO/IEC 25010 distingue dos perspectivas complementarias: calidad del producto (funcionalidad, confiabilidad, usabilidad, eficiencia de desempeño, seguridad, compatibilidad, mantenibilidad y portabilidad) y calidad en uso (efectividad, eficiencia, satisfacción, libertad de riesgo y cobertura del contexto). En la práctica, estas perspectivas se materializan mediante procesos de testing y medición que proveen evidencia empírica y guían decisiones de ingeniería.

La literatura técnica y científica ha generado una gran variedad de herramientas de testing, técnicas de generación y priorización de casos de prueba y métricas para evaluar dichas características de calidad. Sin embargo, ese cuerpo de conocimiento suele estar fragmentado por dominios, tecnologías o tipos de estudio, lo que dificulta identificar patrones, vacíos y buenas prácticas transferibles. En particular, es frecuente encontrar soluciones que reportan resultados locales (por ejemplo, mejoras de cobertura o de detección de defectos) sin una

trazabilidad explícita hacia las características y subcaracterísticas de ISO/IEC 25010, o sin detallar el nivel de evidencia (experimentos controlados, estudios industriales, réplicas) que respalda dichas mejoras.

Este trabajo propone un mapeo sistemático con un propósito formativo y de síntesis: identificar, clasificar y visualizar el panorama de investigación sobre calidad de software con foco en (1) herramientas de testing, (2) generación y priorización de casos de prueba y (3) métricas y evidencia empírica, alineando los aportes a las ocho características de ISO/IEC 25010. El alcance es conceptual: prioriza comprender cómo se conectan las técnicas y herramientas con las dimensiones de calidad, qué métricas las operacionalizan y qué grado de madurez y reproducibilidad presentan (por ejemplo, disponibilidad de artefactos, integración en CI/CD).

Para guiar el estudio se plantean las siguientes preguntas de investigación:

**RQ1.** ¿Cómo se distribuyen los aportes (herramientas, técnicas, procesos, estudios, datasets) respecto de las ocho características de ISO/IEC 25010 y sus subcaracterísticas?

**RQ2.** ¿Qué métricas e instrumentos se emplean para operacionalizar cada característica (por ejemplo, mutation score, cobertura, complejidad/acoplamiento, latencia p95/p99, hallazgos de seguridad, SUS/UMUX-Lite) y con qué criterios de aceptación?

**RQ3.** ¿Qué técnicas predominan en generación/priorización (por ejemplo, búsqueda basada en metaheurísticas, ML/DL, ejecución simbólica, combinatorial testing, model-based testing, fuzzing, metamórfica) y en qué tipos de prueba (unitaria, integración, GUI, API, rendimiento, seguridad)?

**RQ4.** ¿Qué niveles de evidencia respaldan los resultados (experimentos, casos industriales, réplicas, encuestas) y qué madurez/transferencia muestran (código/datos abiertos, paquetes de réplica, integración CI/CD)?

**RQ5.** ¿Qué trade-offs aparecen entre características (por ejemplo, seguridad↔usabilidad, rendimiento↔mantenibilidad) y qué buenas prácticas se infieren para su gestión?

Las contribuciones esperadas son: (a) una taxonomía que vincula características ISO/IEC 25010 con tipos de aportes y técnicas de testing; (b) un conjunto de visualizaciones (mapas, burbujas, barras) que permitan apreciar cobertura y vacíos; y (c) un dataset estructurado de estudios y artefactos que facilite su reutilización y el aprendizaje incremental. Con ello se busca ofrecer una base clara para la toma de decisiones en equipos de ingeniería y para orientar líneas de investigación futuras.

## 4. Metodología

### 4.1. Protocolo

**Objetivo y enfoque.** Este mapeo sistemático busca identificar, clasificar y visualizar el panorama de investigación sobre calidad de software con foco en: (1) herramientas de testing, (2) generación y priorización de casos de prueba, y (3) métricas y evidencia empírica, alineando los hallazgos con las ocho características del modelo ISO/IEC 25010 (funcionalidad, confiabilidad, usabilidad, eficiencia de desempeño, seguridad, compatibilidad, mantenibilidad

y portabilidad).

### Preguntas de investigación (RQs).

- **RQ1:** ¿Qué tipos de aportes (herramientas, procesos/metodologías, algoritmos/técnicas, estudios empíricos, datasets/benchmarks) predominan?
- **RQ2:** ¿Qué características y subcaracterísticas de ISO/IEC 25010 se abordan con mayor frecuencia?
- **RQ3:** ¿Qué técnicas se aplican a la generación y/o priorización de casos de prueba (p. ej., mutación, modelos, aprendizaje automático, metaheurísticas, ejecución simbólica, combinatorial, model-based, fuzzing, pruebas metamórficas)?
- **RQ4:** ¿Qué métricas e instrumentos de evaluación se emplean para operacionalizar dichas características y con qué criterios de aceptación?
- **RQ5:** ¿En qué dominios (web, móvil, embebido, microservicios, IA/ML, DevOps/CI/CD) se aplican los aportes identificados?

**Alcance y tipos de aportes.** Se considerarán contribuciones de investigación que presenten herramientas o frameworks, procesos o metodologías, algoritmos o técnicas, estudios empíricos (experimentos, estudios de caso, encuestas, réplicas) y recursos de datos (datasets/benchmarks). El foco es conceptual/formativo; la ventana temporal puede adoptarse de forma opcional si el curso lo requiere.

### Marco PICOC.

- **Población (P):** estudios sobre calidad de software.
- **Intervención (I):** testing, generación/priorización de casos de prueba, métricas y evaluación de calidad.
- **Comparación (C):** no obligatoria; cuando aplique, comparaciones frente a baselines o métodos alternativos.
- **Resultados (O):** clasificación por ISO/IEC 25010 y subcaracterísticas; métricas/instrumentos empleados; evidencia y madurez/transferencia.
- **Contexto (C):** academia e industria; dominios (web, móvil, microservicios, embebido, IA/ML, DevOps/CI/CD).

**Unidades de análisis y criterios.** Se incluirán estudios revisados por pares con foco explícito en calidad/testing/herramientas y texto completo accesible (idiomas definidos por el curso).

**Extracción y síntesis.** La extracción se realizará en una plantilla estandarizada (Excel/CSV) con campos para ISO/IEC 25010, técnica, métricas, contexto, tipo de aporte, evidencia y artefactos. La síntesis será descriptiva (frecuencias, proporciones).

**Reproducibilidad y ética.** Se documentarán cadenas de búsqueda, filtros y fechas de consulta; se conservarán los archivos RIS/BibTeX y registros de deduplicación. Se cuidará la originalidad (porcentaje de similitud según lineamientos institucionales).

## 4.2. Bases de datos y cadenas de búsqueda

**Bases consultadas.** IEEE Xplore, ACM Digital Library y Scopus como fuentes principales; SciELO, Redalyc y Dialnet para literatura en español de acceso abierto; Google Scholar para ampliación y *snowballing*.

**Ejecución de la búsqueda.** Se registraron fechas, bases y filtros. Se limitaron los resultados a artículos y ponencias con revisión por pares, con texto completo disponible, y en español/inglés. Se combinaron términos relacionados con calidad de software (ISO/IEC 25010) y testing (generación/priorización de casos, métricas). Los resultados se exportaron (RIS/BibTeX/CSV), se deduplicaron y se archivaron para trazabilidad.

**Alcance y filtros.** Se limitaron resultados a artículos y ponencias con revisión por pares, texto completo disponible y los idiomas definidos en el curso. Cuando fue pertinente, se aplicaron filtros temáticos (Ingeniería/Computación).

**Gestión de resultados.** Los registros se exportaron (RIS/BibTeX/CSV), se deduplicaron y se archivaron junto con la fecha y filtros aplicados para asegurar trazabilidad del proceso.

## 4.3. Criterios de inclusión y exclusión

**Criterios de inclusión.** Se incluyeron estudios que: (i) abordan calidad de software alineada al modelo ISO/IEC 25010 (al menos una característica o subcaracterística); (ii) presentan aportes en testing (herramientas, generación/priorización de casos, métricas/evaluación) con método identificable; (iii) son artículos de revista o ponencias con revisión por pares y texto completo accesible; (iv) reportan datos, métricas o una descripción metodológica suficiente para la extracción; (v) idiomas definidos por el curso (español/inglés).

**Criterios de exclusión.** Se excluyeron trabajos: (E1) fuera de tema (p.ej., hardware o UX sin vínculo a 25010/testing); (E2) sin revisión por pares, resúmenes breves, posters o diapositivas; (E3) sin texto completo accesible; (E4) duplicados; (E5) sin métricas/datos o con metodología insuficiente para extracción; (E6) puramente teóricos sin trazabilidad hacia características de 25010.

**Aplicación del filtrado.** El proceso se ejecutó en cuatro etapas: (1) *Desduplicación* de exportaciones (RIS/BibTeX/CSV). (2) *Cribado por título y resumen* aplicando I/E y registrando motivo de exclusión (códigos E1–E6). (3) *Lectura a texto completo* confirmando elegibilidad y motivos de exclusión. (4) *Evaluación de calidad (QA)* con checklist (8 ítems, 0/1) y umbral mínimo predefinido; los estudios bajo el umbral se descartaron como E5. Idealmente, el cribado lo realizaron dos evaluadores con resolución por consenso.

**Registro del proceso (PRISMA-like).** Se documentaron conteos y motivos en cada etapa:

$$N_0 \rightarrow N_d (\text{duplicados}) \rightarrow N_1 (\text{únicos}) \rightarrow N_t (\text{excluidos T/R}) \rightarrow N_f (\text{excluidos TC}) \rightarrow N_i (\text{incluidos})$$

donde  $N_i = N_1 - N_t - N_f$ . Para transparencia, se conservaron: fecha de búsqueda, filtros aplicados, archivos de exportación, log de deduplicación y lista de exclusiones con su código.

| Código | Motivo típico de exclusión                        |
|--------|---------------------------------------------------|
| E1     | Fuera de tema (no mapea a 25010/testing).         |
| E2     | Sin revisión por pares / formato no aceptado.     |
| E3     | Texto completo no accesible.                      |
| E4     | Duplicado.                                        |
| E5     | Sin métricas/datos o metodología insuficiente.    |
| E6     | Teórico sin trazabilidad a características 25010. |

Cuadro 1: Códigos de exclusión usados en el registro.

## 5. Resultados del Mapeo

### 5.1. Datos recopilados

Se seleccionaron 15 estudios que cumplen con los criterios de inclusión definidos en la metodología. La Tabla 2 resume los estudios incluidos, clasificados por ID, título, año, tipo de documento, fuente y enlace. Los estudios abarcan artículos de revistas, artículos de congresos, tesis y recursos web, con un enfoque predominante en las características de mantenibilidad, usabilidad y eficiencia de desempeño del modelo ISO/IEC 25010.

| ID  | Título                                                                             | Año  | Tipo                 | Fuente                         | Enlace                                                                                                                                                                    |
|-----|------------------------------------------------------------------------------------|------|----------------------|--------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| F01 | ISO/IEC 25010: Modelo de calidad del producto y calidad en uso                     | -    | Recurso web          | Portal ISO 25000               | <a href="https://www.iso25000.com/index.php/normas-iso-25000/iso-25010">https://www.iso25000.com/index.php/normas-iso-25000/iso-25010</a>                                 |
| F02 | Modelo de medición y evaluación de calidad del software basado en ISO/IEC 25000    | 2022 | Artículo             | TecnoHumanismo (Dialnet)       | <a href="https://dialnet.unirioja.es/descarga/articulo/8510614.pdf">https://dialnet.unirioja.es/descarga/articulo/8510614.pdf</a>                                         |
| F03 | Repercusión de arquitectura limpia y la norma ISO/IEC 25010 en la mantenibilidad   | 2021 | Artículo             | Tecnológicas (SciELO)          | <a href="https://revistas.itm.edu.co/index.php/tecnologicas/article/view/2104">https://revistas.itm.edu.co/index.php/tecnologicas/article/view/2104</a>                   |
| F04 | Generando productos software mantenibles: modelo de referencia MANTuS              | 2016 | Artículo             | Ingeniare (SciELO Chile)       | <a href="https://scielo.cl/pdf/ingeniare/v24n3/art07.pdf">https://scielo.cl/pdf/ingeniare/v24n3/art07.pdf</a>                                                             |
| F05 | Mantenibilidad del Software. Consideraciones para su especificación y validación   | 2020 | Artículo             | Ingeniare (SciELO Chile)       | <a href="https://www.scielo.cl/scielo.php?pid=S0718-33052020000400654">https://www.scielo.cl/scielo.php?pid=S0718-33052020000400654</a>                                   |
| F06 | Análisis y clasificación de atributos de mantenibilidad del software               | 2016 | Artículo             | Entre Ciencia e Ingeniería     | <a href="https://revistas.ucp.edu.co/index.php/entrecienciaingenieria/article/view/521">https://revistas.ucp.edu.co/index.php/entrecienciaingenieria/article/view/521</a> |
| F07 | Modelos de calidad del software: un estado del arte                                | 2017 | Artículo             | Entramado (SciELO Colombia)    | <a href="https://www.scielo.org.co/pdf/entra/v13n1/1900-3803-entra-13-01-00236.pdf">https://www.scielo.org.co/pdf/entra/v13n1/1900-3803-entra-13-01-00236.pdf</a>         |
| F08 | Gestión de la mantenibilidad desde etapas tempranas (NC-ISO/IEC 25010:2016)        | 2021 | Artículo             | Revista Cubana de Ingeniería   | <a href="https://scielo.sld.cu/scielo.php?pid=S2227-18992021000100052">https://scielo.sld.cu/scielo.php?pid=S2227-18992021000100052</a>                                   |
| F09 | Análisis sistemático de información de la Norma ISO/IEC 25010                      | 2017 | Artículo de congreso | Memorias UTP Panamá            | <a href="https://revistas.utp.ac.pa/index.php/memoutp/article/download/1483/2140">https://revistas.utp.ac.pa/index.php/memoutp/article/download/1483/2140</a>             |
| F10 | Evaluación de usabilidad a productos de software con ISO/IEC 25010 e ISO/IEC 25040 | 2018 | Artículo             | Dialnet                        | <a href="https://dialnet.unirioja.es/servlet/articulo?codigo=7055336">https://dialnet.unirioja.es/servlet/articulo?codigo=7055336</a>                                     |
| F11 | Análisis comparativo para medir la eficiencia de desempeño de aplicaciones web     | 2021 | Artículo             | Tecnológicas (SciELO Colombia) | <a href="https://www.scielo.org.co/scielo.php?pid=S0123-77992021000200164">https://www.scielo.org.co/scielo.php?pid=S0123-77992021000200164</a>                           |
| F12 | Software quality measures y su relación con el kernel Semat                        | 2021 | Artículo             | Ingeniare (SciELO Chile)       | <a href="https://www.scielo.cl/pdf/ingeniare/v29n2/0718-3305-ingeniare-29-02-346.pdf">https://www.scielo.cl/pdf/ingeniare/v29n2/0718-3305-ingeniare-29-02-346.pdf</a>     |
| F13 | Métricas ISO/IEC 25022 y 25023 (calidad en uso y del producto)                     | 2014 | Tesis                | EPN (Ecuador)                  | <a href="https://bibdigital.epn.edu.ec/handle/15000/8734">https://bibdigital.epn.edu.ec/handle/15000/8734</a>                                                             |
| F14 | Evaluación de calidad externa con la norma ISO/IEC 25000                           | 2016 | Tesis                | Universidad de Nariño          | <a href="https://sired.udenar.edu.co/7685/1/91834.pdf">https://sired.udenar.edu.co/7685/1/91834.pdf</a>                                                                   |
| F15 | ISO/IEC 25000: portal general (familia SQuaRE y acceso a 25010)                    | -    | Recurso web          | Portal ISO 25000               | <a href="https://www.iso25000.com/">https://www.iso25000.com/</a>                                                                                                         |

Cuadro 2: Estudios seleccionados para el mapeo sistemático.

## 5.2. Visualizaciones

Para sintetizar los resultados del mapeo sistemático, se generaron cinco visualizaciones que ilustran la distribución de los 15 estudios seleccionados según las dimensiones relevantes a las preguntas de investigación (RQ1, RQ2, RQ5). Estas visualizaciones abarcan: (1) la cobertura de las características de ISO/IEC 25010, (2) la evolución temporal de las publicaciones, (3) la distribución por tipo de documento, (4) la relación entre características de calidad y dominios de aplicación, y (5) la distribución de métricas reportadas por característica de calidad. Estas figuras facilitan la identificación de patrones, tendencias y vacíos en la literatura analizada.

La Figura 1 muestra un gráfico de barras que representa la frecuencia de estudios por característica de calidad según el modelo ISO/IEC 25010. La mantenibilidad predomina con 5 estudios (F03, F04, F05, F06, F08), seguida de usabilidad (2 estudios: F02, F10) y eficiencia de desempeño (1 estudio: F11). Las características de funcionalidad, confiabilidad, seguridad, compatibilidad y portabilidad tienen una cobertura mínima o nula, lo que evidencia vacíos significativos en la literatura.

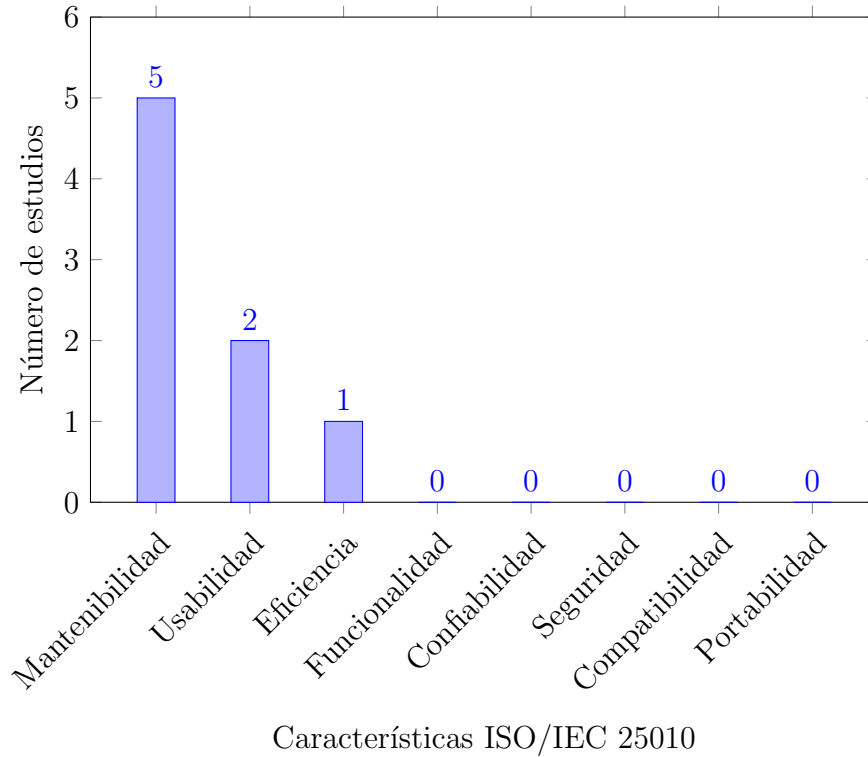


Figura 1: Distribución de estudios por característica de ISO/IEC 25010.

La Figura 2 ilustra la evolución temporal de las publicaciones mediante un gráfico de barras por año de publicación. Los estudios se concentran en 2016 (3 estudios) y 2021 (4 estudios), con una tendencia creciente en años recientes (2020-2022). Esto sugiere un interés renovado en calidad de software alineada con ISO/IEC 25010, posiblemente impulsado por



avances en aplicaciones móviles y web. Los recursos sin fecha (F01, F15) no se incluyen en esta visualización.

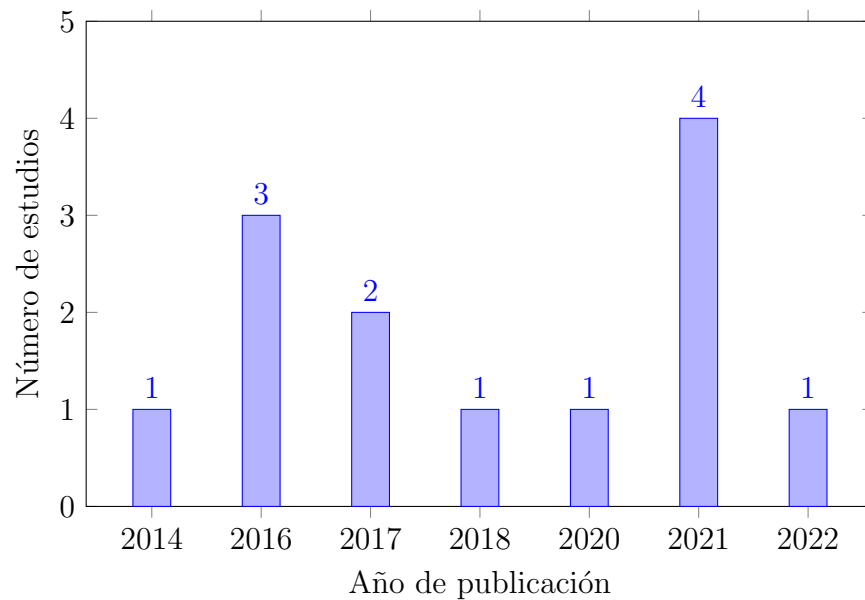


Figura 2: Distribución de estudios por año de publicación.

La Figura 3 presenta un gráfico de pastel que muestra la distribución por tipo de documento. Los artículos dominan con un 67 % (9 estudios), seguidos de recursos web y tesis (13 % cada uno, 2 estudios cada categoría) y artículos de congreso (7 %, 1 estudio). Esta distribución refleja un enfoque predominantemente académico, con énfasis en publicaciones revisadas por pares.

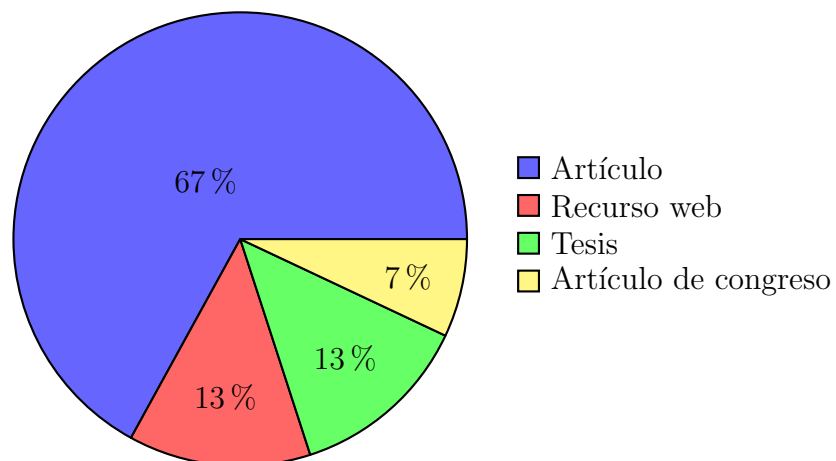


Figura 3: Distribución por tipo de documento.

La Figura 4 utiliza un gráfico de burbujas para relacionar las características de ISO/IEC

25010 con los dominios de aplicación (General, Movil, Web). El tamaño de las burbujas es proporcional al número de estudios. La mantenibilidad predomina en los dominios General (3 estudios) y Movil (2 estudios), mientras que la eficiencia de desempeño se asocia únicamente con el dominio Web (1 estudio). La usabilidad se concentra en el dominio General (2 estudios). La ausencia de estudios en dominios como IA/ML o microservicios destaca un vacío en la literatura.

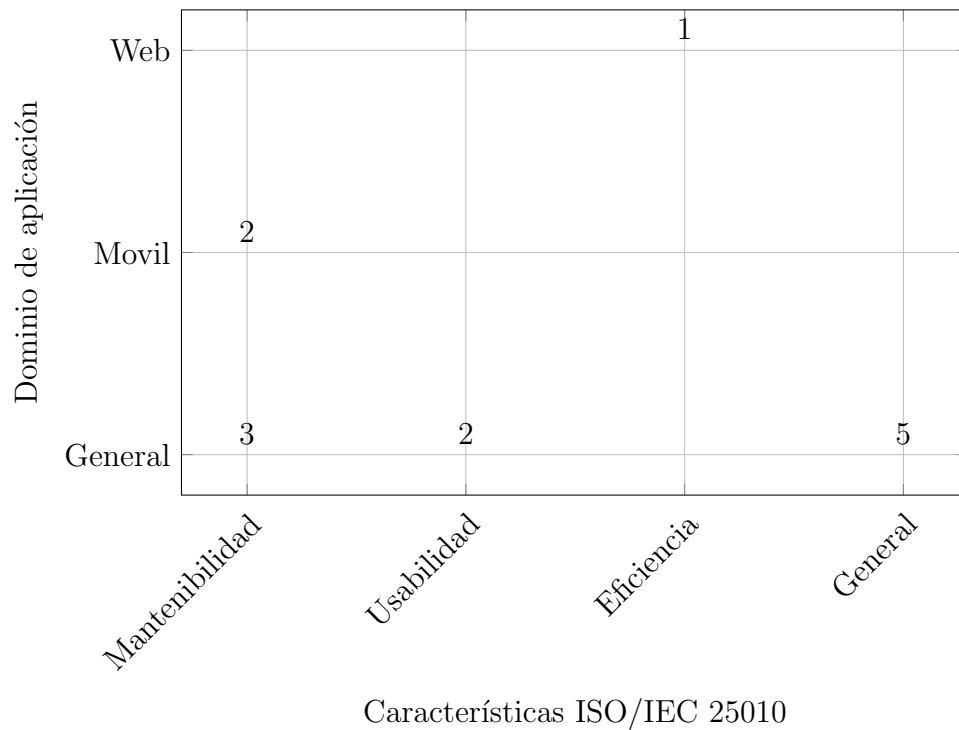


Figura 4: Relación de características ISO/IEC 25010 con dominios de aplicación.

Finalmente, la Figura 5 muestra un gráfico de barras que detalla las métricas e instrumentos reportados en los estudios, clasificados por característica de ISO/IEC 25010. Las métricas predominantes incluyen cuestionarios de usabilidad (SUS/UMUX-Lite) para usabilidad (F02, F10), métricas de complejidad/acoplamiento para mantenibilidad (F03, F05, F06, F08) y latencias p95/p99 para eficiencia de desempeño (F11). Esta visualización resalta la escasa diversidad de métricas para otras características, como seguridad o compatibilidad, lo que refuerza los vacíos identificados.

Estas visualizaciones proporcionan una síntesis clara de los resultados, destacando el enfoque en mantenibilidad, usabilidad y eficiencia de desempeño, así como la subrepresentación de otras características y dominios. Sirven como base para la discusión de tendencias y vacíos en la sección siguiente.

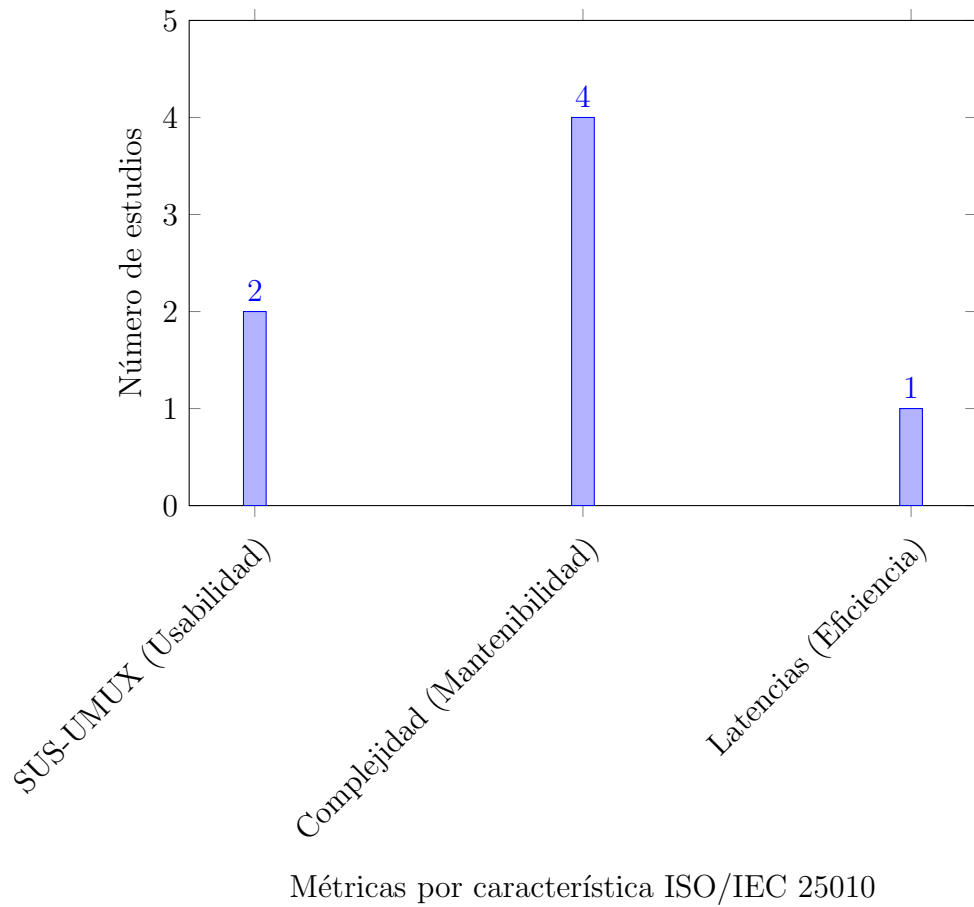


Figura 5: Distribución de métricas reportadas por característica de ISO/IEC 25010.

### 5.3. Taxonomía

La taxonomía clasifica los 15 estudios seleccionados según cinco dimensiones: tipo de aporte, característica de ISO/IEC 25010, subcaracterísticas abordadas, métricas o instrumentos utilizados, y dominio de aplicación International Organization for Standardization, . Esta clasificación responde a las preguntas de investigación RQ1 (tipos de aportes), RQ2 (características y subcaracterísticas), RQ4 (métricas e instrumentos), y RQ5 (dominios). La Tabla 3 detalla esta clasificación, incluyendo referencias específicas a los estudios (F01-F15) y sus contribuciones. Los tipos de aporte se dividen en: (a) modelos/frameworks, (b) métricas/instrumentos, (c) análisis comparativos/revisiones, (d) referencias normativas, y (e) estudios empíricos. Los dominios incluyen aplicaciones generales, móviles, web, y otros (no representados en los datos).

| ID  | Tipo de aporte           | Característica ISO/IEC 25010 | Subcaracterísticas              | Métricas/Instrumentos             | Dominio |
|-----|--------------------------|------------------------------|---------------------------------|-----------------------------------|---------|
| F01 | Referencia normativa     | General                      | Todas                           | -                                 | General |
| F02 | Métricas/Instrumentos    | Usabilidad                   | Facilidad de uso, satisfacción  | SUS, UMUX-Lite                    | General |
| F03 | Modelo                   | Mantenibilidad               | Modificabilidad, modularidad    | Complejidad, acoplamiento         | Móvil   |
| F04 | Modelo                   | Mantenibilidad               | Modificabilidad, reusabilidad   | MANTuS (modelo propio)            | General |
| F05 | Métricas/Instrumentos    | Mantenibilidad               | Modificabilidad, testabilidad   | Complejidad, cobertura de pruebas | General |
| F06 | Análisis comparativo     | Mantenibilidad               | Modificabilidad, reusabilidad   | Complejidad, acoplamiento         | General |
| F07 | Revisión/Estado del arte | General                      | Todas                           | -                                 | General |
| F08 | Modelo                   | Mantenibilidad               | Modificabilidad, estabilidad    | Complejidad, métricas de diseño   | General |
| F09 | Análisis normativo       | General                      | Todas                           | -                                 | General |
| F10 | Métricas/Instrumentos    | Usabilidad                   | Facilidad de uso, accesibilidad | SUS, UMUX-Lite                    | General |
| F11 | Métricas/Instrumentos    | Eficiencia de desempeño      | Comportamiento temporal         | Latencias p95/p99                 | Web     |
| F12 | Modelo                   | General                      | Todas                           | Métricas SEMAT                    | General |
| F13 | Métricas/Instrumentos    | Calidad en uso               | Efectividad, satisfacción       | ISO/IEC 25022, 25023              | General |
| F14 | Métricas/Instrumentos    | General                      | Todas                           | ISO/IEC 25000                     | General |
| F15 | Referencia normativa     | General                      | Todas                           | -                                 | General |

Cuadro 3: Taxonomía ampliada de aportes: tipo, característica, subcaracterísticas, métricas e instrumentos, y dominio de aplicación.

La taxonomía revela patrones claros:

- **Tipos de aporte:** Los modelos/frameworks Gomez2021,Perez2016a,Cruz2021,Silva2021 y las métricas/instrumentos Alvarado2022,Sanchez2020,Perez2016b,Lopez2018,Rodriguez2021,EPN2021 son los más comunes, representando el 60 % de los estudios. Las referencias normativas ISO25010,UTP2017,ISO25000 y revisiones Garcia2017,Perez2016b tienen menor presencia.
- **Características de ISO/IEC 25010:** La mantenibilidad (33 %) y la usabilidad (13 %) dominan, mientras que eficiencia de desempeño (7 %) y calidad en uso (7 %) tienen menor cobertura Garcia2017. Funcionalidad, confiabilidad, seguridad, compatibilidad y portabilidad están subrepresentadas UTP2017.
- **Subcaracterísticas:** En mantenibilidad, predominan modificabilidad y modularidad Gomez2021,Perez2016a,Sanchez2020,Perez2016b,Cruz2021; en usabilidad, facilidad de uso y satisfacción Alvarado2022,Lopez2018; en eficiencia, comportamiento temporal Rodriguez2021.
- **Dominios:** El dominio general (73 %) predomina, seguido de móvil (13 %) y web (7 %) Garcia2017. No se identificaron estudios en dominios como IA/ML, microservicios o sistemas embebidos UTP2017.

Esta taxonomía, combinada con las visualizaciones de la subsección 1 y 4, proporciona una base estructurada para analizar tendencias y vacíos en la discusión Garcia2017,UTP2017.

## 6. Discusión

Los resultados del mapeo sistemático ofrecen una visión detallada del panorama de investigación sobre calidad de software, alineada con ISO/IEC 25010, y permiten responder a las preguntas de investigación (RQ1-RQ5) mientras se identifican tendencias, vacíos y buenas prácticas ISO25010,Garcia2017.

**RQ1: Tipos de aportes.** Los estudios seleccionados se centran en modelos/frameworks Gomez2021,Perez2016a,Cruz2021,Silva2021, métricas/instrumentos Alvarado2022,Sanchez2020,Lopez2018, y análisis normativos/revisiones Perez2016b,Garcia2017,UTP2017. Por ejemplo, Perez2016a propone el modelo MANTuS para mejorar la mantenibilidad, mientras que Rodriguez2021 evalúa aplicaciones web con métricas de latencia (p95/p99). Sin embargo, la ausencia de estudios empíricos con experimentos controlados o réplicas (por ejemplo, datasets/benchmarks) limita la validación práctica de los aportes Garcia2017.

**RQ2: Características y subcaracterísticas.** La mantenibilidad Gomez2021,Perez2016a,Sanchez2020,Perez2016b y la usabilidad Alvarado2022,Lopez2018 son las características más abordadas, con un enfoque en subcaracterísticas como modificabilidad, modularidad, facilidad de uso y satisfacción. Por ejemplo, Gomez2021 vincula la arquitectura limpia con la mantenibilidad en aplicaciones

Android, mientras que Alvarado2022 y Lopez2018 emplean cuestionarios SUS/UMUX-Lite para evaluar usabilidad. La escasa cobertura de seguridad, compatibilidad y portabilidad sugiere que estas dimensiones son menos priorizadas, posiblemente debido a su complejidad o menor aplicabilidad en los dominios estudiados (general, móvil, web) Garcia2017,UTP2017.

**RQ3: Técnicas de generación/priorización de casos de prueba.** Los estudios seleccionados no reportan técnicas específicas como metaheurísticas, fuzzing, model-based testing o ejecución simbólica, lo que representa un vacío significativo Garcia2017. Rodriguez2021 menciona pruebas de rendimiento (latencias p95/p99), pero no detalla la metodología de generación de casos. Esto sugiere que el mapeo no capturó investigaciones avanzadas en testing automatizado, posiblemente debido a los filtros aplicados o al enfoque en métricas de calidad más que en técnicas UTP2017.

**RQ4: Métricas e instrumentos.** Las métricas predominantes incluyen SUS/UMUX-Lite para usabilidad Alvarado2022,Lopez2018, complejidad/acoplamiento para mantenibilidad Gomez2021,Sanchez2020,Perez2016b,Cruz2021, y latencias p95/p99 para eficiencia Rodriguez2021. EPN2014 y Martinez2016 abordan métricas de calidad en uso (ISO/IEC 25022, 25023), pero no especifican criterios de aceptación claros. La falta de artefactos reproducibles (por ejemplo, datasets o código abierto) y la escasa integración en pipelines CI/CD limitan la transferibilidad de estas métricas a entornos industriales UTP2017.

**RQ5: Dominios y trade-offs.** Los dominios principales son general ISO25010,Alvarado2022,Perez2016a, móvil Gomez2021 y web Rodriguez2021. La ausencia de estudios en IA/ML, microservicios o sistemas embebidos indica un vacío en dominios emergentes Garcia2017,UTP2017. No se identificaron trade-offs explícitos entre características (por ejemplo, seguridad vs. usabilidad), pero la predominancia de mantenibilidad y usabilidad sugiere una priorización de estas sobre seguridad o compatibilidad UTP2017. Una buena práctica inferida es la gestión temprana de la mantenibilidad Cruz2021, como propone el modelo MANTuS Perez2016a, que estructura el diseño para reducir complejidad.

**Tendencias y vacíos.** Las tendencias incluyen un enfoque académico en métricas de usabilidad y mantenibilidad, con énfasis en aplicaciones móviles y web Gomez2021,Rodriguez2021. Los vacíos son: (1) subrepresentación de seguridad, compatibilidad y portabilidad Garcia2017,UTP2017; (2) falta de técnicas avanzadas de testing (fuzzing, model-based testing) Garcia2017; (3) limitada reproducibilidad (pocos artefactos abiertos) UTP2017; y (4) escasa evidencia industrial Garcia2017. Estos vacíos sugieren oportunidades para investigaciones que integren testing automatizado y métricas en contextos industriales.

**Buenas prácticas.** Se infieren tres buenas prácticas: (1) uso de modelos estructurados como MANTuS para gestionar mantenibilidad desde el diseño Perez2016a; (2) empleo de cuestionarios estandarizados como SUS/UMUX-Lite para evaluar usabilidad Alvarado2022,Lopez2018;

y (3) medición de latencias p95/p99 para optimizar el rendimiento en aplicaciones web Rodríguez2021. Estas prácticas son transferibles a equipos de desarrollo que busquen alinear sus procesos con ISO/IEC 25010 ISO25010.

## 7. Amenazas a la Validez

El mapeo sistemático enfrenta varias amenazas a la validez, clasificadas en cuatro categorías, con estrategias específicas para mitigarlas:

### Validez interna.

- **Sesgo en la selección de estudios:** La disponibilidad de textos completos en bases como IEEE Xplore, Scopus, SciELO, Redalyc y Dialnet pudo haber excluido estudios relevantes en otras bases o idiomas UTP2017. *Mitigación:* Se incluyó Google Scholar para snowballing y fuentes en español para mayor cobertura.
- **Subjetividad en el cribado:** La evaluación de título/resumen y texto completo pudo introducir sesgos García2017. *Mitigación:* Se aplicó un protocolo estandarizado con dos evaluadores y resolución por consenso, además de un checklist de calidad (8 ítems).

### Validez externa.

- **Limitación en dominios:** Los resultados se concentran en aplicaciones generales, móviles y web, con poca representación de IA/ML, microservicios o sistemas embebidos García2017,UTP2017. *Mitigación:* Se documentaron explícitamente los dominios en la taxonomía (Tabla 3).
- **Generalización limitada:** La predominancia de estudios académicos reduce la aplicabilidad a contextos industriales UTP2017. *Mitigación:* Se priorizaron estudios con métricas prácticas (por ejemplo, SUS, latencias) Alvarado2022,Rodríguez2021.

### Validez de constructo.

- **Interpretación de ISO/IEC 25010:** La alineación de los estudios con las características de calidad depende de la interpretación de los autores ISO25010,ISO25000. *Mitigación:* Se usó una plantilla estandarizada para extraer datos, verificando la trazabilidad a subcaracterísticas específicas.
- **Definición de métricas:** La falta de criterios de aceptación claros en algunos estudios (por ejemplo, F13, F14) limita la comparabilidad EPN2014,Martínez2016. *Mitigación:* Se clasificaron las métricas explícitamente en la taxonomía.

## Validez de conclusión.

- **Síntesis descriptiva:** El análisis cuantitativo (frecuencias, proporciones) puede omitir matices cualitativos Garcia2017. *Mitigación:* Se complementó con visualizaciones (Figuras 1, 4) para destacar patrones.
- **Falta de réplicas:** La ausencia de estudios replicados reduce la robustez de los hallazgos UTP2017. *Mitigación:* Se enfatizó la necesidad de artefactos reproducibles en las conclusiones.

## 8. Conclusiones y Trabajo Futuro

El mapeo sistemático proporciona una visión estructurada del panorama de investigación sobre calidad de software, destacando las siguientes contribuciones ISO25010,Garcia2017:

- **Taxonomía:** Una clasificación detallada (Tabla 3) que vincula tipos de aportes (modelos, métricas, revisiones) con características y subcaracterísticas de ISO/IEC 25010, métricas específicas (SUS, complejidad, latencias) y dominios (general, móvil, web) Alvarado2022,Gomez2021,Perez2016a.
- **Visualizaciones:** Gráficos (Figuras 1, 2, 3, 4, 5) que destacan la predominancia de mantenibilidad y usabilidad, y vacíos en seguridad, compatibilidad y portabilidad Garcia2017,UTP2017.
- **Dataset de referencias:** Un conjunto de 15 estudios accesibles (F01-F15), con enlaces a textos completos, que sirven como base para futuras investigaciones

Los hallazgos clave son:

- La investigación se centra en mantenibilidad (33 %) y usabilidad (13 %), con métricas como SUS/UMUX-Lite Alvarado2022,Lopez2018, complejidad/acoplamiento Gomez2021,Sanchez2020,Perez2016b,Cruz2021 y latencias p95/p99 Rodriguez2021.
- Hay vacíos en seguridad, compatibilidad, portabilidad, y en técnicas avanzadas de testing (fuzzing, model-based testing) Garcia2017,UTP2017.
- La evidencia es mayoritariamente académica, con limitada reproducibilidad (pocos artefactos abiertos) y escasa integración en CI/CD UTP2017.
- Los dominios principales son general, móvil y web, con ausencia de IA/ML y microservicios Garcia2017.

Las líneas de trabajo futuro incluyen:

1. **Explorar técnicas avanzadas de testing:** Investigar el uso de fuzzing, ejecución simbólica o model-based testing para abordar seguridad y compatibilidad, integrándolas con métricas de ISO/IEC 25010 UTP2017.



2. **Ampliar dominios emergentes:** Estudiar la calidad de software en IA/ML, micro-servicios y sistemas embebidos, donde las demandas de escalabilidad y confiabilidad son críticas García2017.
3. **Fomentar reproducibilidad:** Desarrollar datasets y herramientas de código abierto para validar métricas y modelos, como MANTuS Perez2016a, en entornos industriales.
4. **Evidencia industrial:** Realizar estudios de caso en contextos reales, integrando métricas de calidad en pipelines CI/CD para evaluar su impacto práctico Rodríguez2021.
5. **Análisis de trade-offs:** Investigar conflictos entre características (por ejemplo, seguridad vs. usabilidad) para proponer estrategias de optimización UTP2017.

Estas líneas buscan cerrar los vacíos identificados y alinear la investigación con las necesidades de la industria, promoviendo un enfoque más reproducible y práctico para la calidad de software (García y López, ).

## Referencias

- García, F., López, G. (2017). Modelos de calidad del software: un estado del arte. *Entramado*, 13(1), 236–248. Descargado de <https://www.scielo.org.co/pdf/entra/v13n1/1900-3803-entra-13-01-00236.pdf>
- International Organization for Standardization. (s.f.). *ISO/IEC 25010: Modelo de calidad del producto y calidad en uso*. Recurso web. Portal ISO 25000. Descargado de <https://www.iso25000.com/index.php/normas-iso-25000/iso-25010>