



Desambiguación de reportes de aplicaciones móviles en mantenimiento: propuesta de un vocabulario controlado

Disambiguation of mobile application maintenance reports: proposal of a controlled vocabulary

Alionuska Velázquez Cintra^{1,*} <https://orcid.org/0000-0003-2127-8362>
Juan Pedro Febles Rodríguez² <https://orcid.org/0000-0003-3088-3564>
Ailyn Febles Estrada³ <https://orcid.org/0000-0002-5742-9719>

¹Unión de Informáticos de Cuba, La Habana, Cuba

²Universidad de las Ciencias Informáticas, UCI, La Habana Cuba

³Ministerio de las Comunicaciones, MINCOM, La Habana, Cuba

*Autor para la correspondencia: avelazquez@uci.cu

RESUMEN

El análisis de reportes operativos en aplicaciones móviles enfrenta retos de coherencia y precisión del análisis debido a la ambigüedad del lenguaje natural, lo que dificulta la toma de decisiones en la gestión del mantenimiento. Este artículo propone establecer un vocabulario estandarizado para facilitar la interpretación técnica de los datos. Tras comparar vocabularios controlados, tesauros y ontologías, se selecciona el vocabulario controlado como la solución más ágil y colaborativa. La metodología empleada incluye la identificación de términos recurrentes, definiciones consensuadas de forma interdisciplinaria y un proceso de validación iterativo. Este enfoque optimiza la extracción de información, la detección de anomalías y la identificación de patrones en grandes volúmenes de datos textuales. Como resultado, se favorece una toma de decisiones precisa, alineada con los estándares de calidad y las expectativas operativas de los equipos de ingeniería.

Palabras clave: análisis de opiniones; ambigüedad lingüística; codificación del conocimiento; vocabulario controlado; aplicaciones móviles, equipos de ingeniería.

ABSTRACT

The analysis of user-generated operational reports in mobile apps is often hindered by natural language ambiguity, compromising data accuracy for maintenance management. This article aims to establish a standardized vocabulary to ensure a shared technical interpretation. After evaluating controlled vocabularies, thesauri, and ontologies, the study proposes a controlled vocabulary as a practical and collaborative solution to reduce semantic ambiguity. The methodology involves identifying recurrent terms, reaching interdisciplinary consensus on definitions, and

undergoing iterative validation. This approach enhances information extraction, anomaly detection, and pattern recognition within large textual datasets. Ultimately, implementing this structured language framework optimizes decision-making processes, ensuring they are more accurate and better aligned with both operational requirements and development objectives in multidisciplinary engineering environments.

Keywords: opinion analysis; linguistic ambiguity; knowledge codification, controlled vocabulary, mobile applications, engineering teams.

Recibido: 06/01/26

Entregado: 06/01/26

Introducción

La Ingeniería del Software contemporánea ha evolucionado hacia una redefinición del mantenimiento, dejando de ser una fase correctiva post-entrega para convertirse en un proceso continuo de optimización, esencial para la calidad y sostenibilidad de las soluciones tecnológicas [1]. En el ecosistema de los dispositivos móviles, esta etapa adquiere una complejidad superior debido a la fragmentación de plataformas y el dinamismo de las actualizaciones, lo que exige prácticas de desarrollo y soporte más robustas y adaptativas [2]. Este auge de las aplicaciones móviles ha derivado en la generación masiva de datos textuales provenientes de reportes de operación y mantenimiento.

Dichas descripciones, ricas en información cualitativa, constituyen una fuente invaluable para diagnosticar el estado de los activos e identificar no conformidades. En este contexto, la gestión del conocimiento se posiciona como un pilar de la confiabilidad operativa [3]. Bajo esta premisa, la gestión del mantenimiento se asume como el conjunto de acciones técnicas y administrativas aplicadas para asegurar que los activos cumplan su función, mientras que la mejora continua se define como el esfuerzo sistemático por optimizar los flujos de trabajo y reducir la variabilidad operativa. El enfoque en aplicaciones móviles utilizadas en entornos industriales se caracteriza por la criticidad operativa de los reportes de mantenimiento, donde la ambigüedad semántica no solo afecta la experiencia del usuario, sino que puede impactar la disponibilidad de activos, la continuidad operativa y la toma de decisiones técnicas. A diferencia de otros dominios, en el contexto industrial los reportes ambiguos representan un riesgo operativo y no únicamente comunicacional. Ante este escenario, los equipos de ingeniería y gestión de activos en el entorno de desarrollo de aplicaciones móviles en Cuba se enfrentan al reto de procesar estos volúmenes masivos de información mediante técnicas automatizadas para perfeccionar la planificación del mantenimiento y elevar la eficiencia operativa del software. En este sentido, el procesamiento de grandes volúmenes de datos en entornos industriales permite una analítica predictiva que transforma los reportes cualitativos en indicadores de rendimiento operativo [4],

facilitando una toma de decisiones basada en evidencia y no solo en intuiciones técnicas.

Sin embargo, el lenguaje natural es inherentemente ambiguo. Un mismo término puede tener múltiples significados (polisemia) o distintos términos pueden referirse al mismo concepto (sinonimia), fenómenos que representan un desafío crítico tanto para la comunicación humana como para los sistemas de procesamiento de información [5]. Esta ambigüedad se agrava en el contexto de los reportes operativos, donde resultó ser recurrente el uso de jerga específica, abreviaciones, errores tipográficos y expresiones coloquiales. Por ejemplo, la palabra "lento" puede referirse al rendimiento de un equipo ("la bomba es muy lenta al presurizar"), al servicio al cliente ("la respuesta del soporte es lenta") o incluso a una característica específica ("la animación es lenta" - este último ejemplo podría reorientarse a "la respuesta del sensor es lenta" para mayor coherencia industrial). Una interacción como "se cierra" puede significar que la aplicación se cierra inesperadamente (un crash) o que un componente mecánico "se cierra" o "bloquea" inesperadamente. Esta variabilidad contextual y la ambigüedad terminológica representan un obstáculo significativo para la coherencia en el análisis. Bajo esta premisa, el impacto del Procesamiento de Lenguaje Natural (PNL) en la era de la inteligencia artificial busca precisamente desentrañar estas complejidades, desarrollando modelos que simulen la comprensión humana para reducir la brecha entre el lenguaje coloquial y la interpretación técnica [6]. Si cada miembro de un equipo operativo o de ingeniería interpreta los términos de manera diferente, los resultados del análisis serán inconsistentes, lo que llevó a conclusiones erróneas y decisiones subóptimas en la gestión del mantenimiento y la mejora continua de procesos. Los antecedentes recopilados en nuestro diagnóstico (Fase II) confirman este impacto: la falta de un lenguaje común ralentizó el proceso de análisis (tiempos de clasificación de 15 minutos vs. 9.7 con vocabulario estandarizado), los debates improductivos dieron lugar a la reclasificación manual por desacuerdos. Estos hallazgos se alinean con estudios previos que reportan pérdidas de productividad del 25-40% en equipos multidisciplinarios por ambigüedad terminológica.

Para abordar esta problemática, se hace imperativa la creación de un diccionario de términos especializado que permita la desambiguación efectiva de los reportes operativos. Este artículo analizará los diferentes modelos de diccionarios de términos -vocabulario controlado, tesauro y ontología- [4], evaluando sus fortalezas y debilidades para la resolución de ambigüedades en el contexto industrial. Posteriormente, se propondrá el diseño e implementación de un vocabulario controlado como una solución práctica y eficaz para estandarizar la terminología en el análisis de reportes de incidencias de aplicaciones móviles en entornos industriales, promoviendo la coherencia y mejorando la eficiencia del equipo multidisciplinario. La importancia de este estudio radica en su potencial para optimizar la calidad de los datos para la toma de decisiones en procesos ingenieriles de la industria [5], particularmente en la mejora de la confiabilidad y disponibilidad de los activos. Si bien la metodología propuesta es conceptualmente independiente del idioma, la validación empírica presentada en este estudio se realizó

DESAMBIGUACIÓN DE REPORTES DE APLICACIONES MÓVILES EN MANTENIMIENTO: PROUESTA DE UN VOCABULARIO CONTROLADO

exclusivamente sobre reportes en idioma español, por lo que los ejemplos y resultados cuantitativos deben interpretarse dentro de ese contexto lingüístico.

Métodos

La metodología empleada en esta investigación se estructuró en cuatro fases interconectadas, adoptando un enfoque mixto que combinó la revisión teórica con la validación empírica. Este diseño busca la aplicabilidad práctica en el contexto de la industria de software cubano.

Fase I. Análisis comparativo y selección del modelo

Se estudió y comparó modelos terminológicos (vocabulario controlado, tesauros y ontologías) para seleccionar la solución más adecuada, concluyendo que el vocabulario controlado ofrece el mejor equilibrio entre capacidad de desambiguación, complejidad de implementación y mantenimiento.

Fase II. Diagnóstico de ambigüedad y necesidades del dominio

Se realizó un análisis de datos primarios consistentes en reportes históricos del proceso de gestión de incidencias, recolectados en proyectos de desarrollo de aplicaciones móviles de la Universidad de las Ciencias Informáticas. Se realizó un análisis empírico de 129 reportes históricos de mantenimiento y entrevistas a 10 gestores de expertos (gestores de incidencias con más de 5 años de experiencia). Se utilizaron técnicas como TF-IDF (Term Frequency-Inverse Document Frequency) para discriminar términos con alta relevancia semántica frente a ruido lingüístico [7], y el Método Delphi para consensuar y desambiguar términos problemáticos [8].

Fase III. Diseño del vocabulario controlado

Se estructuró y formalizó el vocabulario, incluyendo términos preferidos y no preferidos, contextos de uso y relaciones jerárquicas y asociativas.

Fase IV. Proceso de gestión y mantenimiento

Se estableció un flujo de trabajo iterativo para la adición, modificación y validación de términos, con actualizaciones trimestrales.

Resultados

Fase I. Análisis comparativo y selección del modelo

La gestión de la terminología es fundamental en cualquier campo especializado, y más aún en disciplinas emergentes basadas en análisis de opiniones (texto) en contextos operativos, donde la polisemia y la sinonimia pueden llevar a interpretaciones erróneas. Los sistemas de representación del conocimiento han evolucionado para abordar estas complejidades, permitiendo la transición desde listas simples de términos hacia estructuras relacionales complejas que garantizan la consistencia semántica [9]. Esta evolución se enmarca en los modernos Servicios de Organización del Conocimiento en Red (NKOS), los cuales representan un marco

general para la interoperabilidad terminológica y la alineación de términos técnicos bajo estándares globales [10]. Estos modelos ofrecen diversos niveles de sofisticación para la resolución de ambigüedades. Las ontologías, por ejemplo, poseen una alta capacidad para resolver estas imprecisiones mediante lógica formal y razonamiento automático [11]. No obstante, su construcción y mantenimiento resultan complejos. Esta característica las hace inviables para equipos que priorizan la agilidad en entornos de transformación digital acelerada [1]. La selección de herramientas para el desarrollo de estos modelos varía significativamente en términos de esfuerzo y especialización técnica necesaria [12], lo que justifica un análisis previo de las capacidades de cada modelo (tabla 1).

Tabla 1 - Comparación de modelos de diccionarios de términos para resolución de ambigüedades

Característica	Vocabulario controlado (listas de términos)	Tesoro	Ontología
Definición	Una lista normalizada y predefinida de términos aceptables para un dominio específico. Simplifica la terminología y mejora la gestión de metadatos [13, 14, 15]	Un vocabulario controlado que organiza los términos por relaciones semánticas (sinonimia, antonimia, jerarquía, asociativas) [17].	Una representación formal y explícita de un dominio de conocimiento, definiendo conceptos, propiedades y relaciones lógicas entre ellos [11], aplicada a la estructuración de activos industriales [20].
Estructura	Principalmente lineal o jerárquica simple (taxonomía) [15].	Red de conceptos con relaciones claramente definidas (jerárquicas y no jerárquicas) [17].	Grafos complejos con clases, subclases, propiedades, instancias y axiomas lógicos [19, 20].
Relaciones entre los términos	Jerárquicas básicas (ej. "es un tipo de"), a veces sinónimos preferidos.	Sinonimia: Término Preferido (TP), Usar Por (UP) Jerárquicas: Término Genérico (TG), Término Específico (TE). Asociativas: Término Relacionado (TR). [16].	Clases/Subclases: "es-un", "parte-de". Propiedades: Atributos de los conceptos. Relaciones arbitrarias: Definidas por el dominio Axiomas: Reglas lógicas que definen la coherencia del conocimiento. [11, 18].
Nivel de formalidad	Bajo a medio. Enfoque práctico en la uniformidad [15].	Medio a alto. Estructura más formal para representar relaciones de significado [16, 17].	Muy alto. Basado en lógica formal, permitiendo inferencias y razonamiento automático [11].
Capacidad para resolver ambigüedades	Baja-Media: Principalmente a través de la normalización de términos (ej., "smartphone" vs. "móvil inteligente"). Ayuda a usar el mismo término para el mismo concepto [7].	Media-Alta: Ayuda a la desambiguación léxica al agrupar sinónimos, identificar términos preferidos y mostrar relaciones contextuales. Permite entender que "banco" puede ser "banco (institución financiera)" o "banco (asiento)" [9, 16].	Muy Alta: Permite una desambiguación semántica profunda. Al modelar conceptos y sus propiedades y relaciones lógicas, puede inferir el significado correcto de un término polisémico o ambiguo en un contexto dado, incluso para relaciones complejas [18, 19].

DESAMBIGUACIÓN DE REPORTES DE APLICACIONES MÓVILES EN MANTENIMIENTO: PROUESTA DE UN VOCABULARIO CONTROLADO

Adaptabilidad	Alta. Ofrece una estructura adaptable que equilibra la precisión con la agilidad operativa [14].	Más flexible que una ontología pura, permite añadir y modificar relaciones con cierta facilidad [17].	Menos flexible debido a la complejidad lógica. Cambios pueden tener implicaciones en cascada y requieren procesos complejos de alineación y validación [18].
Complejidad de construcción	Baja a media. Se puede crear con herramientas de hoja de cálculo o bases de datos simples [7, 14].	Media. Requiere herramientas especializadas (software de tesauros) y conocimiento de principios de indexación [12].	Alta. Requiere expertos en modelado de conocimiento, lógicas descriptivas y herramientas de ingeniería de ontologías (ej., Protégé) [12].
Mantenimiento	Relativamente sencillo.	Medio. Requiere revisión periódica de relaciones.	Alto. La adición de nuevos conceptos o relaciones puede ser compleja y requerir validación lógica [18].
Usos Comunes	Control terminológico, indexación de documentos, consistencia en bases de datos [12, 13].	Recuperación de información (ayuda a encontrar documentos con términos relacionados), normalización de lenguaje, desambiguación de términos.	Representación de conocimiento, razonamiento automático, integración de datos heterogéneos, búsqueda inteligente, sistemas de IA [11, 19].

La elección entre estos modelos depende de la complejidad del dominio, los recursos disponibles y la profundidad de la desambiguación requerida. Aunque las ontologías presentan una mayor capacidad teórica para la desambiguación semántica, esta investigación prioriza la viabilidad operativa, la participación interdisciplinaria y la rapidez de adopción en equipos de mantenimiento. En este contexto, las ventajas formales de ontologías y tesauros no compensan su complejidad de diseño, validación y mantenimiento, especialmente en organizaciones con madurez terminológica inicial o intermedia. Para equipos de desarrollo de aplicaciones móviles que inician procesos de estandarización terminológica, el vocabulario controlado constituye una alternativa eficaz al permitir una implementación ágil, una menor demanda de recursos técnicos y una mejora inmediata en la consistencia del lenguaje, sentando así las bases para una comunicación clara y operativamente sostenible.

En el contexto de la gestión de incidencias de un equipo que desarrolla aplicaciones móviles, se justifica el uso de un vocabulario controlado por ofrecer un equilibrio óptimo entre facilidad de implementación y efectividad inmediata en la estandarización del lenguaje interno. Se contrastó esta solución con las limitaciones de los otros modelos, y se comprobó que puede ayudar en la:

- Mejora de la precisión en la anotación y clasificación de datos.
- Facilidad de la comunicación interdisciplinaria.
- Agilidad del proceso de análisis y la toma de decisiones.

Fase II. Diagnóstico de ambigüedad y necesidades del dominio **El escenario**

El desarrollo de aplicaciones móviles no termina con su lanzamiento. El mantenimiento de software es un ciclo continuo [22] y esencial que abarca desde la

corrección de errores críticos hasta la adaptación a nuevos entornos y la implementación de mejoras.

Tipos de mantenimiento [22] que aseguran la funcionalidad y relevancia de las aplicaciones móviles:

- Correctivo: reparar errores o fallos post-lanzamiento. Tiene como insumo principal los errores reportados por los usuarios.
- Adaptativo: ajustar la app a cambios externos (sistema operativo, hardware, normativas) [2].
- Perfectivo: mejorar cualidades de la aplicación (no errores) [21].
- Preventivo: evitar fallos futuros mediante acciones proactivas, a partir del análisis técnico proactivo [3].

El análisis de los principales desafíos asociados a este proceso evidencia que uno de los problemas más frecuentes es la gestión ineficiente de incidencias, particularmente ante errores técnicos, caídas de servidores o cierres inesperados de la aplicación, situación señalada también en estudios recientes [2].

En el entorno de aplicaciones móviles este problema alcanza una mayor dimensión teniendo en cuenta que los errores y fallos en las aplicaciones móviles (bugs) pueden aparecer en cualquier momento y tener un impacto significativo en la experiencia del usuario y en la reputación de la aplicación, lo que a menudo se traduce en la desinstalación masiva. Según reportes de industria, la tolerancia del usuario hacia los cierres inesperados es mínima, lo que traduce los fallos de software en pérdidas reputacionales inmediatas [21].

Una gestión desorganizada lleva a tiempos de inactividad prolongados y a la insatisfacción del cliente.

ITIL constituye un estándar ampliamente adoptado para aportar orden, eficiencia y mejora continua a la gestión de incidencias en aplicaciones móviles, especialmente en entornos empresariales [1]. Este marco de referencia proporciona flujos de trabajo claros para registrar, clasificar y escalar reportes; sin embargo, su énfasis en procesos estructurados no resuelve por sí solo un desafío crítico: la ambigüedad semántica de los mensajes de los usuarios, donde un mismo término puede expresar significados distintos según el contexto. Esta ambigüedad exige la incorporación de herramientas lingüísticas y técnicas de inteligencia artificial orientadas a la estandarización de la comunicación [6].

Análisis empírico de ambigüedades terminológicas en reportes de usuarios

Comprender las dificultades reales asociadas al uso inconsistente de la terminología [7] en los reportes de mantenimiento de aplicaciones móviles requiere una recolección y análisis de datos primarios.

Para identificar las necesidades terminológicas del dominio, se aplicó un enfoque mixto que combinó el análisis estadístico de datos primarios con técnicas de consulta a expertos:

- Análisis de frecuencia y relevancia (TF-IDF): se procesó una muestra de 129 reportes históricos extraídos de proyectos de desarrollo de aplicaciones móviles de la Universidad de las Ciencias Informáticas, los cuales funcionan

DESAMBIGUACIÓN DE REPORTES DE APLICACIONES MÓVILES EN MANTENIMIENTO: PROUESTA DE UN VOCABULARIO CONTROLADO

como entornos de producción real para el ecosistema empresarial cubano. Se utilizó la técnica *Term Frequency-Inverse Document Frequency* (TF-IDF) para discriminar los términos con mayor carga semántica frente al ruido lingüístico, priorizando aquellos con mayor frecuencia en la muestra para su posterior validación por expertos.

- Método Delphi para la desambiguación: los términos identificados fueron sometidos a un proceso de validación mediante el método Delphi, involucrando a 10 gestores de incidencias con más de 5 años de experiencia (expertos). A través de rondas de consulta, se alcanzó un 90% de acuerdo en la definición de las fichas técnicas, los contextos de uso y la asignación de términos preferidos y no preferidos. Para la validación de las definiciones se utilizó un cuestionario estructurado basado en una escala Likert de 5 niveles, donde los 10 expertos evaluaron la claridad, precisión y pertinencia de cada ficha técnica.

Fase III. Diseño Del vocabulario controlado Los fundamentos

Basado en el análisis comparativo realizado, se seleccionó el vocabulario controlado como el modelo más adecuado para abordar la problemática identificada, al ofrecer un equilibrio entre su capacidad para mitigar ambigüedades, la complejidad de su implementación y mantenimiento, y su adaptabilidad a las necesidades de equipos multidisciplinarios con niveles iniciales o intermedios de madurez en la gestión terminológica.

El diseño del vocabulario controlado propuesto incluyó:

- Estructura de las entradas: Definición de los campos esenciales para cada término (término, definición, contexto de uso, sinónimos preferidos, notas).
- Categorización de términos: Agrupación lógica de los términos para facilitar su consulta y gestión.
- Proceso de gestión y mantenimiento: Propuesta de un flujo de trabajo para la adición, modificación y validación de términos, incluyendo roles y responsabilidades.

Aunque la propuesta se centra en un vocabulario controlado, se conceptualizó la posibilidad de una futura evolución hacia un Tesauro. Esta referencia implica la capacidad del modelo para incorporar, en etapas posteriores, relaciones semánticas complejas (jerárquicas, de equivalencia y asociativas). El valor de dejar prevista esta evolución radica en la sostenibilidad de la solución, asegurando que el sistema no quede obsoleto a medida que las necesidades de gestión del conocimiento del equipo crezcan y requieran una estructura de datos más robusta para la automatización del mantenimiento.

La estructura

El diseño del vocabulario controlado se basó en los resultados del consenso alcanzado en la fase anterior. Cada entrada se formalizó mediante fichas técnicas estandarizadas que transforman la subjetividad del reporte en datos procesables

(tabla 4). El diseño se centró en los 12 términos críticos identificados en el diagnóstico incluyendo la definición de campos esenciales para cada entrada: término preferido (TP), que actúa como el identificador único; términos no preferidos (TNP), que agrupan las variantes coloquiales (sinónimos); y el contexto de uso, que, mediante ejemplos reales de los reportes, guía al gestor en la interpretación. Esta decisión se fundamenta en la necesidad de estandarizar la información técnica para reducir el esfuerzo manual en el triaje de errores [23], ya que la calidad de los reportes impacta directamente en la capacidad de reproducción y resolución de los fallos.

Respecto a la inclusión de campos opcionales como Rol del equipo o Relaciones jerárquicas en el instrumento (Tabla 2), se justifica bajo el principio de agilidad operativa. En organizaciones con una madurez terminológica inicial, la obligatoriedad de estos campos podría ralentizar la adopción de la herramienta. No obstante, se mantienen como opcionales para permitir la escalabilidad futura hacia modelos más complejos como tesauros u ontologías a medida que el equipo evolucione en su gestión del conocimiento.

Esta estructura, detallada en la tabla 2, fue validada mediante el método Delphi con un 90% de acuerdo.

Tabla 2 - Estructura de las entradas del vocabulario controlado

Término preferido (TP)	Se establecen formas estandarizadas y únicas para referirse a conceptos específicos
Definición	Explicación concisa y clara del TP en el contexto de opiniones de usuarios
Términos no preferidos (TNP) / Sinónimos	Se asocian variantes o sinónimos comunes utilizados por los usuarios a los términos preferidos, facilitando la remisión al término estandarizado.
Contexto de uso / Ejemplos	Se incluyen frases reales para ilustrar la aplicación de los términos, ayudando a los equipos a interpretar la retroalimentación ambigua en escenarios específicos.
Relación jerárquica (opcional)	Se sugiere la posibilidad de indicar términos genéricos (TG) o específicos (TE) para mejorar la clasificación y organización del vocabulario.
Notas / Consideraciones	Se añaden aclaraciones o advertencias sobre el uso de los términos.
Rol del equipo (opcional)	Se indica qué roles del equipo están relacionados con cada término para facilitar la colaboración.

La arquitectura de implementación y accesibilidad

Complementando el diseño conceptual, se definió una arquitectura de soporte basada en una plataforma centralizada (versión web funcional). Este componente técnico permite que el vocabulario no sea un documento estático, sino una herramienta de consulta ágil donde los gestores de incidencias buscarán términos en tiempo real. La interfaz fue diseñada bajo principios de usabilidad para asegurar que el proceso de desambiguación no añada esfuerzo al flujo de trabajo, integrando motores de búsqueda semántica que vinculan automáticamente los términos no preferidos (TNP) con sus correspondientes definiciones técnicas consensuadas.

Fase IV. Proceso de gestión y mantenimiento

DESAMBIGUACIÓN DE REPORTES DE APLICACIONES MÓVILES EN MANTENIMIENTO: PROPUESTA DE UN VOCABULARIO CONTROLADO

Se estableció un flujo de trabajo iterativo (Fig. 1) diseñado para garantizar la vigencia y escalabilidad del vocabulario controlado. Este proceso no es lineal; se concibe como un ciclo de mejora continua donde la información recolectada durante el mantenimiento diario de las aplicaciones móviles actúa como insumo para actualizar el modelo.

El flujo opera bajo un esquema de retroalimentación constante: los nuevos términos, variantes coloquiales o cambios en el contexto técnico detectados por los gestores de incidencias son registrados y evaluados. Este dinamismo permite que el vocabulario permanezca alineado con la evolución del lenguaje del usuario y las actualizaciones de las plataformas. Las actualizaciones trimestrales consolidan estos cambios mediante un proceso de validación técnica, asegurando que la herramienta mantenga su capacidad de desambiguación y su utilidad como puente entre el lenguaje natural y la interpretación técnica.

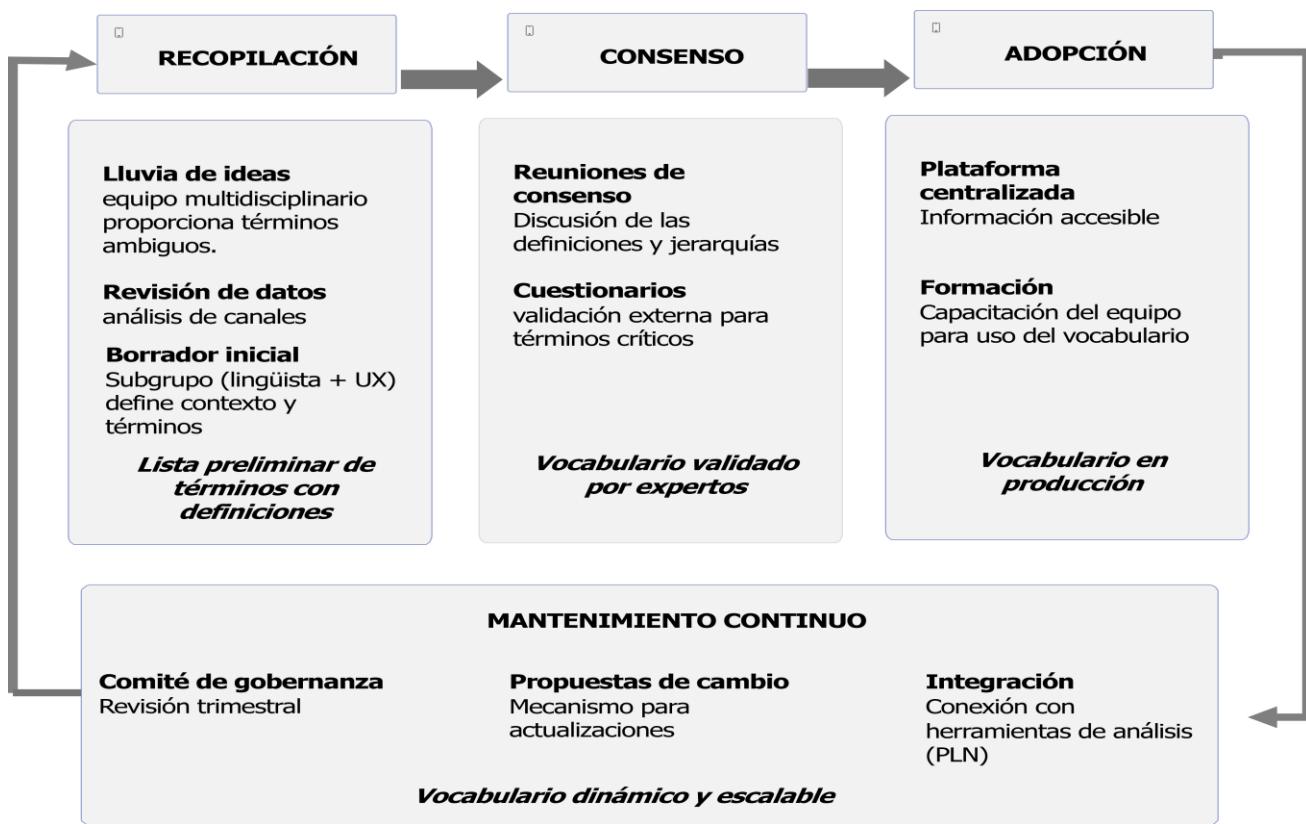


Fig. 1 - Diagrama de flujo del proceso de implementación y gestión del vocabulario controlado (las fases).

Tras la ejecución del piloto (descrito en la metodología), se procedió a comparar el desempeño del equipo bajo los dos escenarios previstos. El procedimiento consistió en el procesamiento paralelo de reportes reales provenientes de los cinco canales de retroalimentación identificados. Se comparó el desempeño del equipo bajo dos escenarios: el método tradicional (basado en interpretación libre del lenguaje natural) y el uso de la arquitectura web centralizada que soporta el vocabulario controlado.

La estrategia de optimización operativa se articuló en cuatro ejes fundamentales:

- Categorización lógica: Agrupación de los 12 términos críticos en 6 categorías para una búsqueda eficiente.
- Precisión en la clasificación: Reducción de la subjetividad mediante definiciones únicas.
- Consistencia interpretativa: Homogeneización del lenguaje entre miembros del equipo.
- Análisis masivo: Mejora en la detección de patrones en grandes volúmenes de datos.

Una vez validado el diseño, el procesamiento paralelo de los 129 reportes permitió contrastar la eficiencia de la propuesta frente al método tradicional. La muestra abarca cinco canales de retroalimentación, lo cual permitió obtener una visión representativa y multifuente de las inconsistencias terminológicas reportadas por los usuarios (tabla 3).

Tabla 3 - Términos ambiguos identificados (por TF-IDF y Delphi)

Término identificado	Frecuencia (TF) / Ocurrencia	Contexto de uso (validado por Delphi)	Ambigüedad detectada (confirmada por experto)	Definición consensuada propuesta (validado por Delphi)	Canales
No funciona	Alta (8)	El usuario reporta una falta general de funcionalidad o respuesta de la aplicación, sin detallar la causa específica.	Alta No especifica el tipo de error: si es de carga, conexión, crash o bloqueo.	La app no responde o no realiza la acción esperada según el usuario, sin una causa técnica clara.	Todos
Error	Alta (6)	El usuario identifica que algo salió mal, generalmente acompañado de un mensaje o un comportamiento inesperado.	Alta Puede ser técnico, de red, de usuario o sin código específico	Situación en que una función genera un fallo técnico identificado o no, a menudo acompañado de un mensaje.	Plataforma de incidencias, Telegram, Facebook
Problemas con la cuenta / problemas con el usuario	Media (5)	El usuario experimenta dificultades para acceder, configurar o gestionar su perfil dentro de la aplicación.	Media Puede referirse a credenciales, bloqueo, errores del sistema o del usuario	Incidencias relacionadas con autenticación, recuperación o gestión de cuenta (inicio de sesión, habilitación de cuenta).	Plataforma, Telegram, Tienda
No se descarga / No puedo descargar	Alta (9)	El usuario no logra obtener un archivo o la aplicación misma, ya sea por problemas de conexión o del sistema.	Alta No se distingue si es por red, error técnico, configuración o compatibilidad	Fallo en el proceso de descarga iniciado desde la app, incluyendo archivos y la propia aplicación.	Todos
Se cayó / Se detiene	Media (4)	El usuario describe un cierre abrupto de la aplicación o que dejó de funcionar repentinamente.	Media Puede referirse a cierre inesperado, desconexión o fallo parcial	Interrupción inesperada del funcionamiento o cierre forzado de la app	Plataforma de incidencias, Facebook
Archivo	Media (5)	El usuario tiene	Media	Incapacidad de	Telegram,

DESAMBIGUACIÓN DE REPORTES DE APLICACIONES MÓVILES EN MANTENIMIENTO: PROUESTA DE UN VOCABULARIO CONTROLADO

multimedia no funciona / no se descarga		dificultades con elementos visuales o de audio, como enviar fotos o reproducir videos.	Puede ser por red, error del sistema, configuración o permisos	enviar, recibir o reproducir archivos multimedia	Facebook
No puedo registrarme / problemas con el registro	Media (6)	El usuario no logra crear una nueva cuenta en la aplicación.	Media No se aclara si es por error del usuario, problema técnico o falta de instrucciones	Imposibilidad de completar el proceso de registro	Todos
Problemas con la instalación	Baja (2)	El usuario tiene dificultades para poner la aplicación en funcionamiento en su dispositivo.	Baja Puede deberse a archivo dañado, incompatibilidad o mala ejecución de pasos	Fallo al intentar instalar la aplicación en el dispositivo	Plataforma de incidencias , App Store, Apkis
No se conecta / no me conecta	Alta (7)	El usuario no logra establecer una conexión con los servicios en línea de la aplicación o acceder a su cuenta.	Alta Puede ser fallo de red, credenciales, bloqueo del usuario o del servidor	La app no establece conexión con el servidor o no permite iniciar sesión	Plataforma de incidencias , Telegram, Facebook
Inconformidad / preocupación / no entiendo	Media (4)	El usuario manifiesta una insatisfacción general o confusión que no necesariamente se deriva de un error técnico específico, sino de la experiencia de uso.	Media Subjetivo, no específico; puede reflejar insatisfacción general	Expresión de molestia o frustración no asociada a un fallo técnico concreto	Telegram, Facebook

El análisis de los reportes históricos reveló que 12 términos clave representaban el 82% de las incidencias ambiguas, con distribuciones significativas ($p < 0.01$) entre canales. La tabla 4 detalla su impacto operativo:

Tabla 4 - Términos ambiguos: frecuencia, impacto y distribución

Término	Frecuencia (TF)	% Incidencias	Nivel ambigüedad	Canales afectados	Interpretaciones distintas
No funciona	8	22%	Alta (4.8/5)	Todos	4
Error	6	18%	Alta (4.5/5)	Plataforma de incidencias, Telegram, Facebook	3
No se conecta	7	15%	Media-Alta (4.2/5)	Plataforma de incidencias, Telegram	2
No se descarga	9	27%	Crítica (4.9/5)	Apkis, App Store, Plataforma de incidencias	5

- El 27% de las incidencias correspondían al término "No se descarga" (TF=9), con 5 interpretaciones técnicas distintas.
- Los términos con ambigüedad "crítica" ($\geq 4.5/5$ en escala Delphi) cubrían el 63% de los casos analizados.
- La dispersión entre canales fue estadísticamente significativa ($\chi^2=12.34$, $*p=0.002*$), siendo Facebook el que mostró mayor variabilidad terminológica.

Hallazgos clave

- Alta frecuencia de términos ambiguos generales: Se detectó que los términos de uso generalizado por los usuarios presentan un índice de ambigüedad elevado al no especificar la causa raíz (ej., carga, conexión, crash, bloqueo). Este fenómeno trasciende lo lingüístico y se manifiesta como un riesgo operativo crítico, ya que dificulta el diagnóstico preciso y genera "tiempos muertos" en la resolución de problemas. Para mitigar este impacto en la eficiencia, los expertos confirmaron estas ambigüedades y propusieron definiciones consensuadas y contextos de uso detallados que permiten transformar el reporte en una unidad de información técnica accionable, permitiendo una toma de decisiones basada en datos estandarizados y no en interpretaciones subjetivas.
- Ambigüedad en temas transversales: La ambigüedad no se limita a fallos técnicos directos, sino que también afecta áreas críticas. En estos casos, la falta de claridad se deriva de múltiples posibles causas subyacentes (credenciales, red, sistema, permisos), lo que subraya la necesidad de una comunicación más estructurada para recopilar información de usuario.
- Variabilidad en canales de retroalimentación: La detección de estos términos ambiguos abarca múltiples canales (Plataforma de incidencias, Telegram, Tienda, Facebook), indicando que la ambigüedad es un problema persistente independientemente de la plataforma de reporte. Esto sugiere la necesidad de implementar estrategias de clarificación de términos a través de todos los puntos de contacto con el usuario. Aunque el problema es persistente en todos los canales, la naturaleza del lenguaje en redes sociales exige que el vocabulario sea lo suficientemente flexible para capturar esa informalidad y traducirla a términos técnicos.
- Relevancia del contexto de uso: La incorporación de un "Contexto de uso" consensuado y validado por expertos es fundamental porque traduce la retroalimentación ambigua a escenarios de usuario más específicos, permitiendo una interpretación uniforme por parte de los equipos de gestión de incidencias y desarrollo.

En conjunto, los hallazgos confirmaron que la alta frecuencia de términos ambiguos y su transversalidad en temas críticos no constituyen únicamente un fenómeno lingüístico, sino un riesgo operativo que afecta directamente la agilidad del mantenimiento al retrasar el diagnóstico y la resolución de incidencias. La variabilidad detectada entre canales —especialmente en entornos de comunicación informal— refuerza la necesidad de un vocabulario controlado que actúe como herramienta puente entre el lenguaje natural del usuario y el lenguaje técnico del equipo de desarrollo. La definición de contextos de uso estandarizados se convierte así en un requisito previo para transformar reportes ambiguos en unidades de información técnica accionable, optimizando los procesos de clasificación de incidencias y fortaleciendo la toma de decisiones conforme a ITIL.

Categorización de los términos

DESAMBIGUACIÓN DE REPORTES DE APLICACIONES MÓVILES EN MANTENIMIENTO: PROPUESTA DE UN VOCABULARIO CONTROLADO

Para optimizar la recuperación de información, los términos se organizaron en 6 categorías principales, empleando estrictamente la nomenclatura definida en la Tabla 2 y la Tabla 4. El diseño se centró en los 12 términos críticos identificados previamente, utilizándolos como cimientos para organizar el conocimiento y derivar de ellos términos específicos y sinónimos relacionados. Este proceso permitió extender la propuesta a un inventario inicial de 45 términos que cubren el espectro operativo más crítico detectado en el diagnóstico. Como se detalla en la tabla 5, esta agrupación lógica permite que, incluso ante el uso de lenguaje coloquial en canales informales, el sistema redirija al usuario hacia una categoría técnica, facilitando una gestión más eficiente donde la Funcionalidad General actúa como el eje central al cubrir el 55% de las incidencias.

Tabla 5 - Categorías principales y ejemplos de términos ambiguos

Categoría	Términos clave incluidos	% de Incidencias cubiertas
Funcionalidad general	no funciona, error, no se conecta	55%
Descargas y actualización	no se descarga, no actualiza	33%
Registro y acceso	problemas con registro, permisos	15%
Rendimiento	lento, se bloquea	15%
Errores graves	crash, pantalla negra	9%
Multimedia	archivo no funciona	10%

Las categorías no son mutuamente excluyentes, ya que una misma incidencia puede involucrar múltiples dimensiones funcionales.

Los resultados obtenidos (ver tabla 6) muestran una optimización de los tiempos y la calidad del proceso. El tiempo promedio de clasificación se redujo en un 35%, logrando bajar de 15 a 9.7 minutos por reporte. Este ahorro es consecuencia directa de la arquitectura de soporte web, que permite una consulta inmediata sin ambigüedades.

Tabla 6 - Métricas de eficacia

Indicador	Antes	Después	Mejora
Tiempo promedio de clasificación por reporte	15 minutos	9.7 minutos	↓35%
Consistencia en interpretación entre equipos	58%	89%	↑31%
Reclasificación manual necesaria	42%	11%	↓74%

La aplicación sistemática de este proceso de consenso elevó la consistencia interpretativa del equipo de un 58% a un 89%, demostrando que la desambiguación mediada por expertos elimina los debates subjetivos durante las reuniones de mantenimiento y mejora la calidad de la información técnica accionable desde el primer contacto. El impacto más significativo se observa en la reclasificación manual, que descendió del 42% al 11%. Esta reducción del 74% en el reproceso demuestra que el vocabulario resuelve la ambigüedad en el primer contacto, evitando que la incidencia circule por múltiples departamentos antes de ser atendida correctamente.

La implementación reveló tres impactos críticos en la gestión:

- Eficacia en la desambiguación: Los términos no funciona y error pasaron de 4 interpretaciones a una única definición técnica. Esto elevó la consistencia del equipo al 89%, eliminando debates subjetivos en las reuniones de mantenimiento.
- Diferencial por canales: La mejora en la plataforma interna (42%) superó a la de redes sociales (28%). Esto confirma que, aunque el vocabulario es una herramienta puente eficaz, la informalidad en Facebook y Telegram sigue siendo el mayor desafío para la extracción de datos limpios.
- Optimización de patrones: La normalización terminológica aceleró en un 20% la detección de fallos recurrentes en grandes volúmenes de texto, facilitando una toma de decisiones proactiva alineada con ITIL.

La reducción del reproceso y de los tiempos de clasificación impacta directamente en indicadores operativos clave del entorno industrial, tales como la disponibilidad del sistema y la eficiencia del mantenimiento (tanto correctivo como preventivo). Los resultados validan el diagnóstico inicial: la ambigüedad crítica en el lenguaje natural de los usuarios (82% de las incidencias) compromete la precisión operativa. No obstante, la implementación del vocabulario controlado mitiga este riesgo, demostrando que un modelo de menor complejidad formal genera mejoras significativas cuando se alinea con las capacidades reales del equipo humano.

A diferencia de marcos como ITIL, que se limitan a estructurar el flujo del proceso, o modelos complejos como las ontologías, la propuesta ofrece un equilibrio óptimo entre rigor técnico y viabilidad operativa. Su diseño práctico y colaborativo, validado por expertos en el entorno industrial cubano, constituye la principal novedad de este estudio: no busca imponer un lenguaje rígido, sino actuar como un traductor eficiente entre el usuario y el ingeniero.

Aunque la investigación se basa en un piloto específico en Cuba, la metodología es generalizable. Se conceptualiza que la madurez de este modelo permitirá, en una etapa futura, evolucionar hacia un tesoro para abordar relaciones semánticas más complejas, manteniendo siempre la optimización de la extracción de información como eje central.

Conclusiones

1. La investigación valida que la ambigüedad en el lenguaje natural (presente en el 82% de las incidencias) compromete la precisión operativa, exigiendo soluciones estandarizadas para garantizar la coherencia en el análisis multidisciplinario. El vocabulario controlado se ratifica como el modelo óptimo por su equilibrio entre eficacia en la desambiguación y facilidad de implementación.
2. El diseño estructural basado en términos preferidos y contextos de uso (Tabla 4) demostró una alta eficacia operativa al elevar la consistencia interpretativa del equipo al 89%. Esto confirma que el modelo no solo es lingüístico, sino una herramienta de gestión que elimina debates subjetivos en el mantenimiento.

3. La implementación piloto evidenció una optimización de los procesos, reduciendo el tiempo de clasificación en un 35% y, fundamentalmente, disminuyendo la reclasificación manual necesaria del 42% al 11%. Esta mejora del 74% en el reproceso asegura que la información técnica sea accionable desde el primer contacto.
4. Se comprobó un comportamiento diferencial según el canal de origen, donde la plataforma interna alcanzó una mejora del 42% frente al 28% en redes sociales. Esto demuestra que el vocabulario controlado funciona como una "herramienta puente" indispensable para traducir la informalidad de canales como Facebook y Telegram hacia categorías técnicas.

Recomendaciones

- Evolucionar el vocabulario controlado hacia una Ontología de Dominio para incorporar relaciones semánticas complejas que favorezcan la comunicación automática entre especialistas y usuarios finales.
- Integrar la arquitectura de soporte web con herramientas de Procesamiento de Lenguaje Natural (PLN) para automatizar la traducción de reportes informales en redes sociales hacia las categorías técnicas definidas.
- Evaluar el impacto de los ciclos de actualización trimestrales en la reducción de la deuda técnica y en la precisión del análisis de términos críticos de alta ambigüedad como "no se descarga" o "no funciona".
- Extender la aplicación del vocabulario controlado a otros idiomas, lo cual requerirá procesos específicos de recolección, consenso y validación terminológica para adaptar el modelo a nuevos contextos lingüísticos.

Referencias

1. Vasilieva, T. (2021). *Digitalization of Industrial Enterprise Processes: Challenges and Opportunities*. HAL Science. <https://hal.science/hal-03264850>
2. Cárdenas-Villavicencio, O. E., et al. (2021). Comparativa de tendencias de desarrollo de software móvil. 3C TIC. Cuadernos de desarrollo aplicados a las TIC, 10(1), 123-145. ISSN: 2254-6529.
3. Cabrera-Hernández, M., León-González, J. L., & Martínez-Casas, J. (2022). La gestión de la información y el conocimiento en el mantenimiento industrial. Revista Universidad y Sociedad, 14(2), 237-246. ISSN: 2218-3620.
4. Ji, S., Wang, J., Xu, W., & Li, Y. (2021). Industrial Big Data Analytics for Smart Manufacturing: A Review. *Journal of Manufacturing Systems*, 60, 204-219. ISSN: 0278-6125.
5. Jurafsky, D., & Martin, J. H. (2024). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition* (3rd ed. draft). Stanford University. <https://web.stanford.edu/~jurafsky/slp3/>

6. Thellmann, S., et al. (2023). Standardizing vocabularies for industrial interoperability. En *ACM International Conference Proceeding Series*. ACM. <https://doi.org/10.1145/3635059.3635104>
7. Rodríguez-Gutiérrez, A. (2023). *Técnicas de minería de texto y sus principales aplicaciones*. [Trabajo de Fin de Máster, Universidad de Cantabria]. Repositorio Institucional UC. <https://repositorio.unican.es/xmlui/handle/10902/30947>
8. Esandi-Baztán, M. Á. (2025). Aplicación del método Delphi en la identificación de unidades especializadas. *Círculo de Lingüística Aplicada a la Comunicación*, 102, 239-260. ISSN: 1576-4737.
9. Pérez-Valera, M. D. (2023). *Contribuciones a la gestión de terminología y la representación del conocimiento*. [Tesis Doctoral, Universidad de Córdoba]. Repositorio Helvia. <https://helvia.uco.es/handle/10396/33024>
10. Hodge, G., Zeng, M. L., & Tudhope, D. (2022). The Networked Knowledge Organization Systems (NKOS) Taxonomy. *Research Ideas and Outcomes*, 8, e94931. ISSN: 2367-7163. <https://doi.org/10.3897/rio.8.e94931>
11. Russell, S. J., & Norvig, P. (2020). *Artificial Intelligence: A Modern Approach*. (4th ed.). Pearson.
12. Alharbi Z, Aldekhail A. A Survey on Ontology and Thesaurus Development Tools. *J Comput Sci Technol*. 2022;37(2):438-452.
13. Ibrahim, M. S. (2024). Metadata Management: Trends and Future Directions. *The Christian Librarian*, 66(1). <https://doi.org/10.5525/tcl.v66i1.4515>
14. Smith, C. H. (2022). Controlled Vocabularies: Past, Present and Future of Subject Access. En *Subject Access: Current Practices and New Trends*. Routledge. ISBN: 9781003211785.
15. Barcelos, P., et al. (2023). A Model-Driven Transformation from Lexicons to Thesauri. [ResearchGate]. DOI: 10.13140/RG.2.2.30232.16645.
16. Chakraborty, S. B. (2024). Online Controlled Vocabulary: A Framework for Knowledge Organization and Dissemination. *International Journal of Information Research and Review*. [ResearchGate].
17. Lund University. (2024). Metadata and Interoperability in Digital Repositories. *Pjos Journal*, 15(1). <https://journals.lub.lu.se/pjos/article/view/26465>
18. Ren S, Zhang S, Li J, et al. A Survey of Ontology Alignment: From Traditional Methods to Deep Learning. *Artif Intell Rev*. 2021;54(1):643-665.
19. Alonso de la Fuente, C. (2023). *Estructuración de información sobre activos industriales mediante el uso de ontologías y LLMs*. [Trabajo de Fin de Máster, Universidad de Cantabria]. Repositorio Institucional UC. <https://repositorio.unican.es/xmlui/handle/10902/30947>
20. MoldStud. Navigating App Maintenance: Common Challenges and Effective Solutions for Schools [Internet]. MoldStud.com; 2024 [citado 2025]

DESAMBIGUACIÓN DE REPORTES DE APLICACIONES MÓVILES EN MANTENIMIENTO: PROUESTA DE UN VOCABULARIO CONTROLADO

- Jul 2]. Disponible en: <https://moldstud.com/articles/p-navigating-app-maintenance-common-challenges-and-effective-solutions-for-schools>
21. Instabug. Mobile User Expectations in 2025: Performance, Privacy, and AI [Internet]. Instabug; 2024. Disponible en: <https://www.instabug.com/mobile-user-expectations-2025>
22. Chaparro, O., Bernal-Cárdenes, C., Lu, J., Moran, K., Marcus, A., Di Penta, M., & Haiduc, S. (2019). Assessing the Quality of the Steps to Reproduce in Bug Reports. En *Proceedings of the 27th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE '19)*. ACM, pp. 86–97. <https://doi.org/10.1145/3338906.3338947>

Conflicto de intereses

Los autores declaran que no hay conflicto de intereses.

Contribución de cada autor:

Alionuska Velázquez Cintra: Fue responsable de la conceptualización y diseño del estudio, definiendo la metodología y los objetivos. Fue responsable directa de la recopilación de los datos experimentales y su análisis formal, así como la interpretación de los resultados. Lideró la redacción del borrador original del manuscrito y coordinó las revisiones.

Juan Pedro Febles Rodríguez: Contribuyó significativamente en la validación de la metodología y el análisis de los resultados, velando por la robustez de los datos. Participó en la revisión crítica y edición del borrador.

Ailyn Febles Estrada: Supervisó la formulación de los objetivos de la investigación, desde la formulación de la hipótesis hasta la validación de las conclusiones. Se aseguró de que el trabajo cumpliera con los estándares académicos y éticos requeridos y verificó su estructura general. Validó la versión final del manuscrito, asegurando su rigor académico.

Sobre los autores

Alionuska Velázquez Cintra: Ingeniera en Ciencias Informáticas. Máster en Calidad de Software con experiencia en planificación, análisis, pruebas y mantenimiento de sistemas de software. Profesor Auxiliar con más de 12 años de experiencia como docente universitaria. Imparte clases de pregrado y postgrado en asignaturas relacionadas con la ingeniería, gestión y calidad de software. Miembro del Board de pruebas de Cuba (CuTB) y posee varias certificaciones internacionales como Tester. Tiene publicado varios artículos en revistas y en ponencias de eventos científicos.

Juan Pedro Febles Rodríguez: Profesor Titular de la Universidad de las Ciencias Informáticas, Cuba, Lic. en Matemáticas, Máster en Informática y Doctor en Ciencias

Técnicas. Tiene más de 40 años de experiencia en la Educación Superior. Es experto en acreditación de Maestrías. Miembro de varios consejos científicos de universidades. Editor asociado y experto de revistas científicas, con amplia experiencia en la evaluación de programas. Es consultor en educación a distancia y gestión del conocimiento.

Ailyn Febles Estrada: Profesora Titular de la UCI y profesora invitada de varias universidades cubanas y extranjeras. Graduada como Ingeniera Informática de la Universidad Tecnológica de La Habana (CUJAE). Máster en Informática Aplicada y Doctora en Ciencias Técnicas. Tiene más de 20 años de experiencia en la Educación Superior. Presidenta del Board de pruebas de Cuba (CuTB) certificada internacionalmente como Tester. Conferencista de eventos nacionales e internacionales, coordinadora de programas y proyectos de innovación asociados a la transformación digital.