

# Evolución de la estimación del esfuerzo en el desarrollo de software: De los modelos tradicionales a la automatización basada en IA y DevOps

Iliana Lizbeth Alvarado Lara<sup>1</sup>, Blanca Dina Valenzuela Robles<sup>1</sup>,  
René Santaolaya Salgado<sup>1</sup>, Gabriel González Serna<sup>1</sup>,  
Noé Alejandro Castro Sánchez<sup>1\*</sup>, Miguel Ángel Hidalgo Reyes<sup>2</sup>

<sup>1</sup>Tecnológico Nacional de México/CENIDET, Ciencias Computacionales,  
México

<sup>2</sup> Instituto Tecnológico Superior de Xalapa/ITSX, Veracruz,  
México

d23ce054@cenidet.tecnm.mx, blanca.vr@cenidet.tecnm.mx, rene.ss@cenidet.tecnm.mx,  
gabriel.gs@cenidet.tecnm.mx, noe.cs@cenidet.tecnm.mx, miguel.hr@xalapa.tecnm.mx

**Resumen.** La estimación del esfuerzo y los recursos es crucial para el éxito en el desarrollo de software, un campo que enfrenta desafíos constantes debido a su complejidad y dinamismo. Este estudio analiza la evolución de las técnicas de estimación del esfuerzo en el desarrollo de software desde la adopción de metodologías ágiles hasta la consolidación del paradigma DevOps, mediante un análisis bibliométrico de 4,196 documentos de Scopus (2008-2024) utilizando SciMAT y VOSviewer. Los resultados evidencian una transición de modelos tradicionales, como COCOMO y puntos de función, hacia enfoques basados en inteligencia artificial, redes neuronales y modelos predictivos, destacando tres áreas clave: gestión de proyectos, técnicas automatizadas y tecnologías avanzadas. Los mapas estratégicos muestran que la estimación del esfuerzo en el desarrollo de software es el tema más influyente, mientras que fiabilidad del software y optimización mantienen su relevancia, y emergen tendencias como algoritmos de aprendizaje automático y árboles de decisión. A nivel global, Estados Unidos, China y Alemania lideran la producción científica. Se identificaron instituciones influyentes dentro de este campo, como la Universidad del Sur de California y la Universidad del Oeste de Canadá, aunque estas no necesariamente pertenecen a los países con mayor volumen de publicaciones, lo que sugiere que su impacto se debe a la calidad y relevancia de sus contribuciones científicas. Finalmente, se destaca la necesidad de desarrollar herramientas híbridas que integren modelos clásicos y técnicas avanzadas, promoviendo enfoques de estimación del esfuerzo en el desarrollo de software en entornos DevOps y ágiles para

mejorar la precisión y eficiencia en la gestión de proyectos de software.

**Palabras clave.** Análisis bibliométrico, estimación de proyectos de software, metodologías ágiles, DevOps, inteligencia artificial.

## Evolution of Effort Estimation in Software Development: From Traditional Models to AI-Based Automation and DevOps

**Abstract.** The estimation of effort and resources is crucial for success in software development, a field that constantly faces challenges due to its complexity and dynamism. This study analyzes the evolution of effort estimation techniques in software development, from the adoption of Agile methodologies to the consolidation of the DevOps paradigm, using a bibliometric analysis of 4,196 Scopus documents (2008-2024) with SciMAT and VOSviewer. The results show a transition from traditional models, such as COCOMO and function points, toward approaches based on artificial intelligence, neural networks, and predictive models, highlighting three key areas: project management, automated techniques, and advanced technologies. Strategic maps indicate that software development effort estimation is the most influential topic, while software reliability and optimization maintain their relevance, and trends like machine learning algorithms and decision trees are emerging. Globally, the United States, China,

and Germany lead scientific production. Influential institutions in this field were identified, such as the University of Southern California and the University of Western Canada, although they do not necessarily belong to the countries with the highest volume of publications, suggesting their impact is due to the quality and relevance of their scientific contributions. Finally, the study highlights the need to develop hybrid tools that integrate classic models and advanced techniques, promoting software development effort estimation approaches in DevOps and Agile environments to improve accuracy and efficiency in software project management.

**Keywords:** Bibliometric analysis, software project estimation, agile methodologies, DevOps, artificial intelligence.

## 1. Introducción

La estimación del esfuerzo en los proyectos de desarrollo de software es un elemento crítico para el éxito empresarial. Desde los primeros días de la informática, prever con precisión los recursos, el tiempo y el alcance de los proyectos ha sido imperativo. En un entorno dinámico y altamente competitivo como la industria del software, donde la rapidez en el lanzamiento de productos, la satisfacción del cliente y la gestión eficaz de recursos son esenciales, la capacidad de realizar estimaciones precisas y realistas se convierte en un pilar fundamental para la competitividad y la sostenibilidad a largo plazo de las empresas de desarrollo de software [14].

El impacto de la estimación del esfuerzo en el desarrollo de software inicial en la planificación del proyecto es considerable, incluso cuando estas estimaciones no se basan en datos históricos o conocimientos expertos [15]. La mayoría de los proyectos enfrentan desafíos relacionados con excesos de esfuerzo y/o incumplimientos de los plazos [17]. A medida que las metodologías de Desarrollo de software han evolucionado desde la década de 1970, han surgido nuevos desafíos y oportunidades en el proceso de estimación [10].

Los estudios bibliométricos representan herramientas fundamentales para analizar la evolución de los enfoques metodológicos y su influencia en distintas áreas de investigación [18, 11, 13, 16]. Particularmente, este estudio bibliométrico tiene impacto en el área de las

estimaciones del esfuerzo en el desarrollo de software [9].

Garfield introdujo el concepto de la *cienciometría* en la década de 1950 y sentó las bases para los estudios bibliométricos [22]. En su reconocido artículo de 1955, propuso un sistema bibliográfico que evita la cita acrílica de datos incompletos u obsoletos [12]. Además, introdujo la cita como una medida cualitativa del impacto académico y promovió conceptos como indexación de citas y enlace de citas, que sirvieron como bases para los motores de búsqueda académica [5].

Siguiendo esta línea, el presente estudio tiene como objetivo analizar la evolución de las estimaciones de esfuerzo en el desarrollo de proyectos de software a través de un enfoque bibliométrico, analizando colaboraciones entre autores, tendencias de investigación y colaboraciones internacionales.

Con el propósito de explorar estos aspectos y proporcionar una visión completa del estudio, el artículo se organiza de la siguiente manera: la sección 2 contiene la revisión de la literatura; la sección 3 detalla la metodología empleada; en la sección 4 se presentan los resultados; la sección 5 ofrece una discusión de los hallazgos; y, finalmente, la sección 6 expone las conclusiones del estudio.

## 2. Revisión de la literatura

Uno de los primeros modelos para realizar estimación del esfuerzo en el desarrollo de software, y ampliamente aceptado, fue COCOMO, desarrollado por Barry Boehm en 1981, el cual utiliza métricas basadas en el tamaño del software para prever costos y esfuerzos necesarios en su desarrollo [7].

Otro enfoque clásico es el uso de puntos por función, propuesto por Allan Albrecht en 1979, que mide el tamaño funcional del software, independientemente del lenguaje de programación o la tecnología utilizada. Este método se enfoca en la funcionalidad desde la perspectiva del usuario, lo cual lo hace aplicable a diferentes plataformas y lenguajes de desarrollo [2].

El método Delphi es otra técnica clásica de estimación que se utiliza para obtener estimaciones más precisas y consensuadas sobre el esfuerzo, tiempo o costos requeridos para desarrollar un proyecto de software. En este método, un grupo de expertos es consultado en múltiples rondas de preguntas, y sus respuestas son revisadas y ajustadas en función de la recopilación y análisis de las opiniones del grupo. Este método, aunque útil, depende en gran medida del juicio subjetivo de los expertos, y puede ser susceptible a sesgos [8].

A lo largo de las últimas décadas, estas técnicas clásicas han constituido la base de la estimación de esfuerzos en proyectos de software [20]. Sin embargo, con la rápida evolución de la tecnología y las metodologías de desarrollo, estas técnicas han sido complementadas o adaptadas para satisfacer las necesidades de proyectos más dinámicos y complejos [19].

Con el surgimiento de metodologías ágiles, como Scrum, la estimación del esfuerzo en el desarrollo de software ha evolucionado hacia enfoques más iterativos y flexibles [24]. La estimación ágil es una de las tendencias actuales, que se basa en realizar estimaciones en cada sprint o iteración, en lugar de prever el esfuerzo de todo el proyecto desde el principio [4].

Técnicas como el planning poker y las historias de usuario permiten realizar estimaciones rápidas y consensuadas en un entorno colaborativo, adaptándose a los cambios frecuentes en los requisitos y en el alcance del proyecto [13].

Otra tendencia moderna es la aplicación del aprendizaje automático en la estimación del desarrollo de proyectos de software, como redes neuronales y árboles de decisión, que se utilizan para identificar patrones en proyectos previos y realizar predicciones basadas en datos históricos [21].

Además, los modelos de aprendizaje profundo, como el deep learning, también han comenzado a aplicarse en las estimaciones del esfuerzo, especialmente en proyectos donde se cuenta con grandes conjuntos de datos [23]. En paralelo, la combinación de técnicas clásicas y modernas ha dado lugar a enfoques híbridos [6]. Por ejemplo, algunos proyectos utilizan COCOMO o puntos por función junto con IA para mejorar la estimación del esfuerzo en el desarrollo de software, logrando un

balance entre la experiencia histórica y el análisis de datos en tiempo real [1].

Estos enfoques reflejan la necesidad de adaptar las metodologías de estimación a los requerimientos cambiantes de la industria del software.

A pesar del progreso en las técnicas de estimación de esfuerzo, existe un vacío en la literatura en cuanto a estudios bibliométricos exhaustivos sobre este tema. La mayoría de las investigaciones se centran en presentar o mejorar técnicas específicas, pero pocas han analizado la evolución de estas técnicas desde una perspectiva cuantitativa.

Además, no se ha investigado en profundidad cómo han evolucionado las redes de colaboración entre autores ni las tendencias en el uso de técnicas de estimación del esfuerzo en el desarrollo de software en diferentes regiones o contextos a lo largo del tiempo.

Otra limitación es la falta de estudios que comparen de manera integral el impacto de técnicas modernas, como la inteligencia artificial, con técnicas tradicionales en términos de precisión, eficiencia y aplicabilidad en diferentes tipos de proyectos. Este vacío de información impide una comprensión global del desarrollo y la eficacia de las metodologías de estimación del esfuerzo en el desarrollo de software.

Este estudio aborda estos vacíos a través de un análisis bibliométrico que permita mostrar la evolución de las técnicas de estimación del esfuerzo en el desarrollo de software, identificar tendencias emergentes y áreas de colaboración, y evaluar el impacto de nuevas tecnologías en comparación con los métodos tradicionales.

### 3. Metodología

El diseño de este estudio es tratado a través de un análisis bibliométrico como método para examinar la evolución de la producción científica sobre las estimaciones del esfuerzo en el desarrollo de proyectos de software durante el periodo 2008-2024.

El proceso metodológico que se siguió para el análisis bibliométrico se muestra en la Fig.1, el cual es el siguiente:

### 3.1. Definir objetivos

En este estudio, se busca identificar el desempeño de los componentes de investigación como las redes de colaboración entre autores las instituciones, países y revistas en el área de las estimaciones del esfuerzo en el desarrollo de proyectos de software.

Asimismo, se busca analizar la estructura de las redes temáticas y los mapas científicos mediante técnicas bibliométricas, para identificar temas emergentes patrones de co-ocurrencia de palabras clave y evaluar la influencia de estos estudios en la evolución del conocimiento.

### 3.2. Identificar fuentes de datos

Para este estudio, se seleccionó la base de datos Scopus por ser una fuente confiable y relevante para la recopilación de datos bibliográficos, debido a su extensa cobertura de literatura científica y su capacidad para proporcionar métricas de citación, coautoría y palabras clave.

Se optó por Scopus en lugar de otras bases de datos como Web of Science o IEEE Xplore debido a su mayor cobertura en ciencias de la computación y software engineering, así como por la disponibilidad de herramientas avanzadas de análisis bibliométrico. Además, Scopus ofrece una integración más efectiva con las herramientas utilizadas en este estudio (SciMAT y VOSviewer), lo que permite un procesamiento de datos más eficiente y estructurado.

### 3.3. Generar cadena de búsqueda

Se genera la cadena de búsqueda que es utilizada para extraer los datos relevantes. Este paso es crucial, ya que la calidad y pertinencia de los resultados dependen de una estrategia de búsqueda bien diseñada. La cadena de búsqueda incluyó términos clave relacionados con el tema de estudio, así como sinónimos y combinaciones de palabras clave que reflejan las áreas de interés definidas en el objetivo del estudio: (software development OR software engineering) AND (estimation OR effort estimation OR cost estimation OR software estimation OR time estimation).

### 3.4. Recopilar datos bibliográficos

Se aplicó la cadena de búsqueda previamente definida en la base de datos de Scopus con filtros específicos para optimizar los resultados: tipo de búsqueda general, obteniendo inicialmente 7,318 documentos; año de publicación (2008 a 2024), reduciendo el conjunto a 5,069 documentos; área temática (ciencias computacionales e ingeniería), limitando a 4,179 documentos; idioma (inglés), resultando en un total de 4,588 documentos; y tipo de documento (artículos y conferencias), quedando finalmente 4,196 documentos seleccionados para el análisis.

Esta consulta recopiló información relevante, incluyendo el título, autores, afiliaciones, año de publicación, palabras clave, resúmenes y referencias de cada documento. Además, se llevó a cabo un proceso de limpieza de datos, eliminando registros duplicados y documentos irrelevantes. La aplicación de estos filtros proporcionó un conjunto de datos confiables y representativos, asegurando que solo se incluyeran publicaciones pertinentes para el estudio.

### 3.5. Identificar herramientas para el análisis

Se seleccionó VOSviewer y SciMAT como herramientas para realizar el análisis de los datos recopilados. Estas herramientas especializadas en la construcción y visualización de redes bibliométricas son ideales para analizar grandes volúmenes de datos de publicaciones científicas. Permiten crear mapas de redes para identificar relaciones y patrones en el conjunto de datos, facilitando el análisis de coautoría, ocurrencia de palabras clave y citación, entre otros.

Las métricas clave implementadas para el análisis bibliométrico son: número de ocurrencias, número de citas, peso total, fuerza de asociación, conteo completo y conteo fraccional, densidad de publicación y cita, y clústeres y agrupamiento.

### 3.6. Seleccionar el tipo de análisis

Se seleccionó el tipo de análisis a realizar para explorar las relaciones y patrones en los datos recopilados. VOSviewer y SciMAT permiten realizar varios tipos de análisis, cada uno enfocado

en un aspecto específico. El análisis de coautoría examina las colaboraciones entre autores, instituciones o países, identificando las redes de colaboración en un campo de estudio. El análisis de ocurrencia de palabras clave mide la frecuencia y relación de palabras clave en los documentos analizados, dando como resultado la evolución de temas durante el periodo.

El análisis de citación evalúa el impacto de documentos, autores, fuentes o países al medir el número de veces que han sido citados. El análisis de co-citación se centra en la frecuencia con la que dos documentos, autores o fuentes son citados conjuntamente, lo que permite identificar temas o enfoques metodológicos relacionados. El análisis de acoplamiento bibliográfico observa las relaciones entre documentos, autores o instituciones que comparten referencias comunes, indicando similitudes temáticas y enfoques de investigación alineados.

### **3.7. Visualizar datos gráficamente**

Se visualizan los datos gráficamente utilizando VOSviewer y SciMAT. Este paso es fundamental, ya que transforma los datos obtenidos en representaciones visuales que facilitan la interpretación y el análisis de patrones, relaciones y tendencias dentro del campo de estudio. Se pueden crear tres tipos de mapas para visualizar los resultados. El primero de ellos es un mapa de red, que muestra las conexiones entre los elementos en forma de red, donde los nodos representan los elementos (como autores o palabras clave) y las líneas que los conectan indican relaciones o co-ocurrencias.

El segundo es un mapa de superposición, el cual añade una dimensión temporal o categórica a la red, permitiendo ver cómo ciertos elementos evolucionan con el tiempo o se agrupan por características específicas, como el año de publicación; y el último es el mapa de densidad, que resalta áreas con alta densidad de elementos, lo que permite identificar las zonas de mayor actividad o interés en el campo de estudio.

### **3.8. Interpretar resultados**

Este paso consiste en analizar y dar sentido a las visualizaciones obtenidas en los pasos previos,

con el fin de interpretar los resultados proporcionando una comprensión integral de la evolución y el estado actual de la investigación en la estimación para el desarrollo de proyectos de software. Este análisis también permite identificar oportunidades para futuras investigaciones y posibles áreas de mejora en el campo.

### **3.9. Síntesis de resultados y conclusiones**

Se presentan de forma estructurada los hallazgos del estudio sobre estimación para el desarrollo de proyectos de software. Destacando el impacto en el campo, tendencias y áreas de oportunidad para futuras investigaciones, consolidando así el aporte del estudio al conocimiento en la estimación para el desarrollo de proyectos de software.

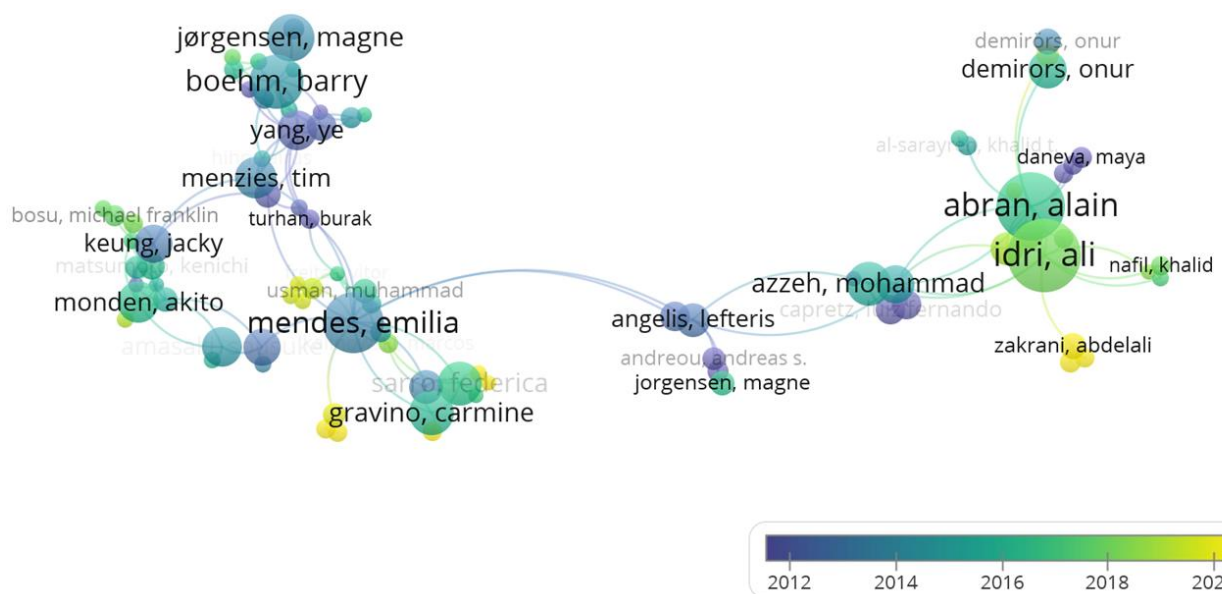
## **4. Resultados**

Esta sección presenta los principales hallazgos obtenidos a través de este análisis bibliométrico. Donde se incluyen enfoques de evaluación, tales como el análisis de coautoría, la evolución de los temas en distintos subperiodos, y el impacto bibliométrico de los clústeres temáticos en términos de centralidad, densidad e índice h. Además, se presenta el análisis de citación y las relaciones entre instituciones y países. Estos resultados ofrecen una base para comprender las tendencias actuales y futuras en la estimación para el desarrollo de software, resaltando tanto las áreas consolidadas como los temas emergentes con potencial de crecimiento.

### **4.1. Análisis del desempeño**

#### **4.1.1. Análisis de coautoría**

La Fig. 2 muestra a los autores clave, sus colaboraciones y su impacto en la investigación sobre estimación para el desarrollo de proyectos de software. La visualización de red en la parte superior izquierda representa las relaciones de coautoría y el impacto de los autores en el campo. Los nodos representan a los autores, y las conexiones indican colaboraciones entre ellos. El tamaño de los nodos refleja la frecuencia de



**Fig. 2.** Redes de colaboración y autores más citados en la investigación sobre estimaciones para el desarrollo del software (2008-2024).

citación de cada autor, mientras que el color muestra la evolución temporal de sus publicaciones. Algunos autores destacados en la red incluyen a Mendes, Boehm, Jørgensen, Yang, Menzies, Idri, Abran, Mohammad, Demirors, y Gravino, quienes tienen un fuerte impacto y múltiples colaboraciones en el campo.

#### 4.1.2. Análisis de citación

La Fig. 3 muestra un análisis de citas por organizaciones en el campo de la investigación, utilizando una visualización de red para indicar la relación y el impacto de las instituciones dentro del área de estimación del esfuerzo en el desarrollo de software. En la red, los nodos representan diferentes organizaciones académicas y corporativas, mientras que las conexiones reflejan la intensidad de colaboración entre ellas.

Las instituciones con mayor impacto en términos de citas incluyen la Universidad del Sur de California, la Universidad del Oeste de Canadá y la empresa NFA Estimation Inc. Se observa que estas instituciones han acumulado un alto número de citas, consolidándose como actores clave en la evolución de este campo. Sin embargo, se aprecia que la mayoría de las colaboraciones se concentran en instituciones de Estados Unidos,

China y Alemania, lo que coincide con los países líderes en producción científica sobre el tema.

La Fig. 4 muestra un análisis de redes de colaboración entre países en el campo de esta investigación, visualizando las relaciones y la intensidad de las colaboraciones científicas entre ellos.

En esta red, los nodos representan a los países, y el tamaño de cada nodo refleja la frecuencia y relevancia de cada país en términos de participación en investigaciones. Las líneas que conectan los nodos indican las colaboraciones, y el grosor de estas líneas refleja la intensidad o frecuencia de colaboración entre países.

Los países influyentes y colaborativos son Estados Unidos (con 14,196 citas), China (con 7,017 citas), Alemania (con 6,260 citas), Reino Unido (con 5,499), e India (con 4,730 citas), los cuales presentan conexiones con otros países, lo que sugiere una red de colaboración científica amplia y diversa. Además, destacan importantes conexiones entre países europeos, como Italia (con 2,871 citas), Francia (con 2,744 citas), España (con 2,035 citas), Noruega (con 2,032 citas), Suecia (con 1,762 citas), República Checa (con 1,425 citas) y Suiza (con 1,087 citas), así como entre países asiáticos, como India (con

**Table 1.** Relación entre centralidad y densidad en los clústeres analizados del subperíodo 2008 - 2013

Clúster	Centralidad	Densidad
Estimation	97.54	19.02
Soft computing Techniques	42.7	30.87
Models	26.26	2.68
Project Managers	25.42	1.97
Forecasting	24.99	6.68
Software Development Projects	24.79	3.29
Artificial Intelligence	24.41	3.61
Software Reliability	22.56	16.98
Function Point	19.22	8.47
COCOMO	7.61	3.60
Information Systems	6.56	1.93
Embedded Systems	5.52	8.94
Software Measurement	4.60	1.99
Functional Size	4.49	10.71
Vectors	3.93	18.59
Requirement Engineering	1.40	4.50
Project Cost Estimation	1.21	5.42
Nonlinear Systems	0.25	17.58

4,730 citas), Hong Kong (con 1,932 citas) y Japón (con 1,478 citas). Algunos países menos destacados, como México (con 246 citas), Bélgica (con 189 citas) y Egipto (con 182 citas), aunque tienen nodos pequeños, muestran conexiones con múltiples países, lo que podría indicar un crecimiento en su participación en redes de colaboración internacional.

#### 4.2. Análisis en términos de ciencia por subperíodo

El análisis se realizó considerando tres subperíodos dentro de la ventana de tiempo establecida (2008-2024), permitiendo identificar tendencias y cambios en cada etapa.

La segmentación de los subperíodos seleccionados responde a criterios cuantitativos y cualitativos que permiten evaluar con mayor precisión la evolución de un campo de estudio. Factores como el volumen de publicaciones,

cambios temáticos, solapamiento de términos y redes de colaboración son determinantes para la segmentación del período. En este sentido, establecer cortes temporales permite comprender cómo han evolucionado las tendencias, qué enfoques han ganado relevancia y cuáles han perdido impacto en las investigaciones.

##### 4.2.1. Análisis del subperíodo 2008-2013

La estimación del esfuerzo en el desarrollo de software ha evolucionado con la aparición de DevOps en 2008 y la consolidación de metodologías ágiles, lo que ha transformado los enfoques tradicionales de estimación.

En este análisis, se empleó un enfoque basado en agrupación de términos a través de coocurrencia de palabras clave. Los clústeres fueron generados mediante la métrica de fuerza de asociación, que mide la relación entre términos dentro de la base de datos seleccionada. Los temas fueron seleccionados con base en su frecuencia de aparición y su relevancia dentro de la red de publicaciones.

Los valores de centralidad y densidad asignados a cada clúster se obtuvieron mediante el análisis de redes bibliométricas, identificando aquellos temas con mayor interconectividad dentro del dominio de estudio.

La centralidad indica el nivel de conexión de un clúster con otros temas dentro de la red bibliométrica. Un valor alto de centralidad sugiere que el clúster es un tema clave en la investigación, actuando como puente entre diferentes áreas del conocimiento.

La densidad, en cambio, mide el grado de desarrollo interno del clúster. Un valor alto de densidad implica que el tema ha sido ampliamente explorado dentro de su propia área, mientras que un valor bajo sugiere que el tema aún está en desarrollo y presenta oportunidades de investigación futura.

La distribución de los clústeres identificados en el subperíodo 2008-2013 junto con sus valores de centralidad y densidad se presentan en la Tabla 1, los cuales permiten caracterizar su relevancia dentro de las estimaciones en el desarrollo de software. Se destaca el clúster de *Estimation*, con un grado de conexión de 97.54 que representa la centralidad, posicionándose como el tema clave en el desarrollo del campo.

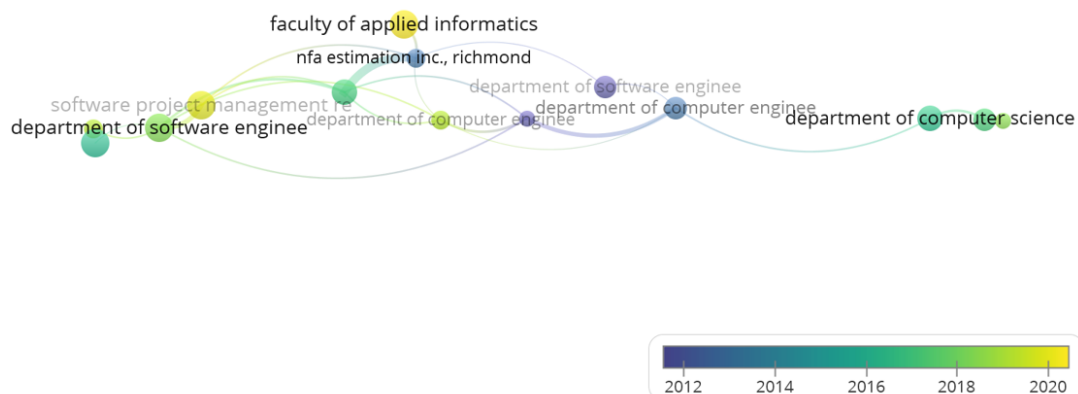


Fig. 3. Citas por instituciones (2008-2024).

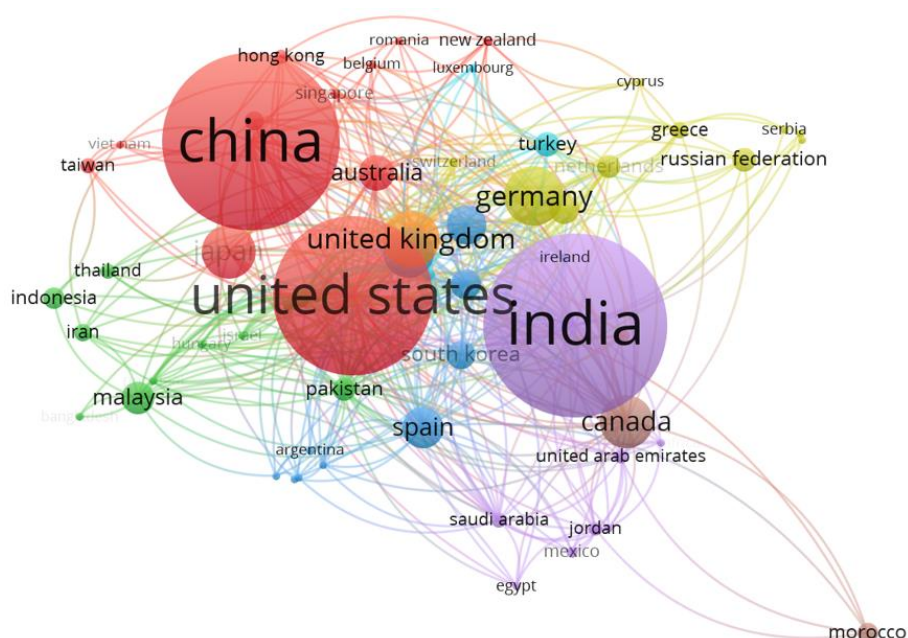


Fig. 4. Citas por países (2008-2024).

Sin embargo, su grado de desarrollo interno moderado, es decir, la densidad, sugiere que aún hay margen para fortalecer su cohesión interna.

Este tema estratégico tiene el potencial de actuar como un puente para integrar otras áreas de conocimiento como estimación del esfuerzo, proyectos de software y estimación de costos, entre otros. Asimismo, las *Softcomputing Techniques* también sobresalen como un tema estratégico durante este período.

Por otro lado, temas como *Artificial Intelligence*, *Models* y *Forecasting* presentan un alto nivel de conexión con otros temas, pero muestran un bajo grado de desarrollo interno.

Esto los convierte en áreas relevantes para futuros trabajos interdisciplinarios que busquen establecer conexiones con otros temas dentro del campo.

En contraste, temas como *Project Managers*, *Information Systems* y *Software Development*



**Table 2.** Propiedades del clúster *Estimation*

Propiedad	Valor
Número total de documentos	415
h-index	44
Total de citas	7,355

*Projects* se posicionan como áreas de menor relevancia en la actualidad, pero con un alto potencial de crecimiento. Estas áreas podrían fortalecerse a través de investigaciones más específicas y cohesionadas, lo que facilitaría su integración en la red temática general.

Finalmente, aunque los clústeres de *Function Point* y *Project Cost Estimation* están bien desarrollados internamente, su baja conexión con otros temas sugiere que se encuentran relativamente aislados. Estas áreas podrían fortalecerse mediante la exploración de aplicaciones prácticas que las vinculen con temas centrales, como la *Artificial Intelligence*.

Las propiedades del clúster principal asociado al tema clave de *Estimation* son descritas en la Tabla 2. Este clúster incluye un total de 415 documentos analizados, los cuales representan los trabajos de investigación vinculados al tema y que fueron considerados en el análisis. Este número no solo refleja el volumen significativo de investigación generada durante el período estudiado, sino que también subraya la relevancia del tema como uno de los principales ejes de interés y atención dentro de la comunidad científica. La magnitud de este conjunto de documentos refuerza la importancia de la estimación del esfuerzo en el desarrollo de software como un área clave de estudio y desarrollo en el ámbito científico.

El indicador de impacto científico (h-index) es una propiedad que mide la calidad de las publicaciones asociadas al tema. En este caso, un h-index de 44 significa que al menos 44 documentos relacionados con el tema han recibido 44 o más citas cada uno. Este valor es considerablemente alto, lo que sugiere que la investigación en *estimación* no solo es extensa, sino también influyente en términos de impacto científico. Además, indica que las publicaciones relacionadas son frecuentemente referenciadas

por otros investigadores, destacando su importancia en el avance del campo.

Por último, el total de citas recibidas, que asciende a 7,355, demuestra el impacto global de este tema en la literatura científica. Este valor subraya la influencia significativa del tema en otros trabajos de investigación como la *Effort Estimation*, *proyectos de software*, *estimación de costos*, *desarrollo de software*, *software para la estimación de costos*, *costos*, *software para la estimación del esfuerzo*, *esfuerzo del desarrollo de software*, *modelos para la estimación* y *software para la estimación*.

Estas propiedades se representan visualmente mediante mapas estratégicos, los cuales ilustran los documentos principales y secundarios asociados a un tema central. En estos mapas, los clústeres temáticos se representan como nodos distribuidos en un cuadrante, donde su posición está determinada por su centralidad (eje X) y densidad (eje Y). La Fig. 5 presenta la distribución de los ejes temáticos del clúster principal *Estimation* con respecto al número total de documentos analizados en el periodo 2008- 2013.

Por lo tanto, el tema de las estimaciones en el desarrollo del software se posiciona como un tema central en el campo de estudio, caracterizado por su alto volumen de publicaciones, su notable calidad científica y su impacto académico considerable, consolidándolo como un eje clave en la red temática.

Los clústeres ubicados en el cuadrante superior derecho corresponden a los temas motores, que son considerados áreas clave en el campo de investigación. Estos temas combinan un alto nivel de conexión con otros clústeres (alta centralidad) y un fuerte desarrollo interno (alta densidad), lo que los posiciona como el núcleo del campo de estudio. Dentro de este cuadrante, el clúster de *Estimation* sobresale como el más destacado, tanto por su posición estratégica como por el tamaño de su nodo, evidenciando su relevancia central en el desarrollo del campo.

El cuadrante superior izquierdo representa los temas altamente desarrollados pero aislados, caracterizados por su baja centralidad y alta densidad. Estas áreas son específicas y maduras, pero representan poca conexión con otros temas dentro del campo de estudio, en el caso de *Embedded Systems* o *Nonlinear Systems*, existen

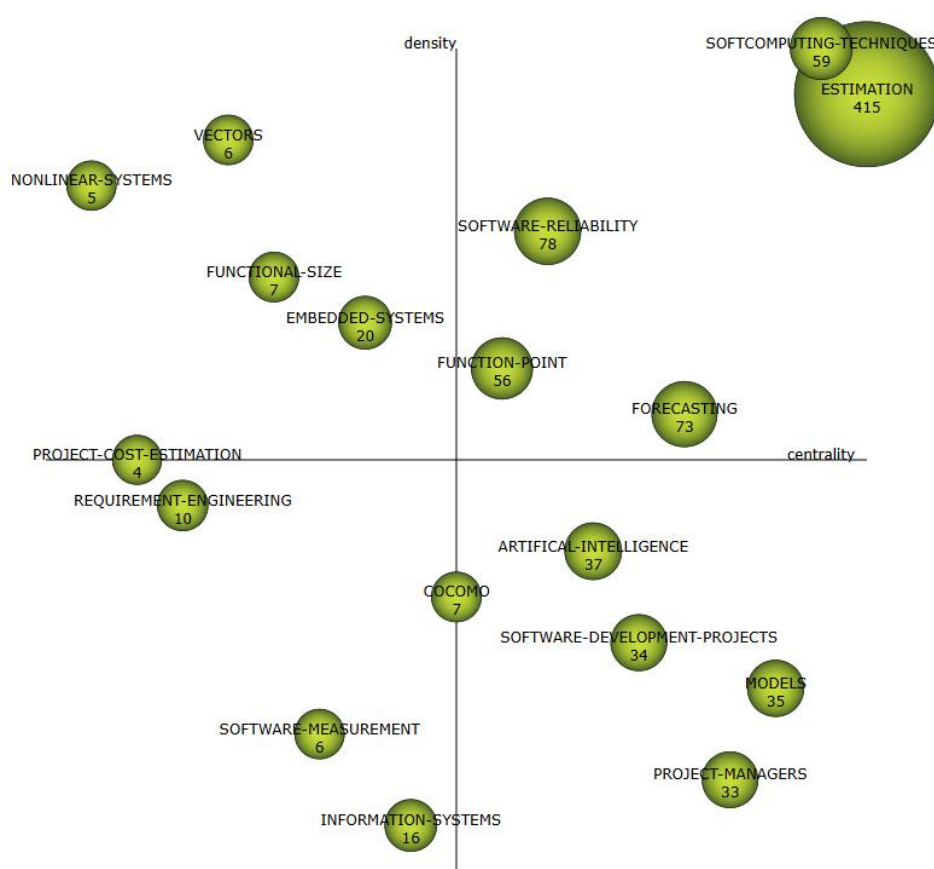


Fig. 5. Mapa estratégico del clúster *Estimation* basado en centralidad y densidad y número total de documentos

teorías, metodologías y aplicaciones muy avanzadas y especializadas que han sido ampliamente exploradas y validadas dentro de su propio ámbito.

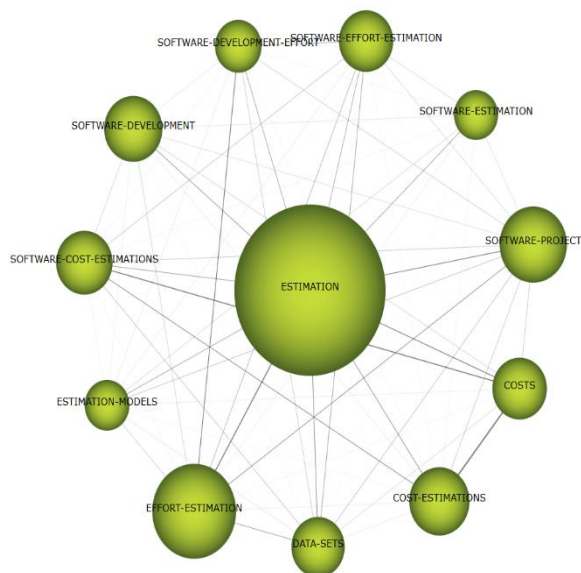
En el caso de *Project Cost Estimation* puede ser un tema muy desarrollado dentro de la gestión de proyectos, pero su aplicación y relevancia están limitadas a contextos específicos, lo que reduce su interacción con otros campos como la inteligencia artificial, *Vectors* es un concepto matemático bien establecido que se aplica en contextos muy específicos, como la ingeniería, pero no suele expandirse hacia áreas emergentes o menos relacionadas.

*Functional Size* es un concepto utilizado principalmente en la medición de software, un área muy especializada dentro de la ingeniería de sistemas. Su aplicación está tan acotada que no suele extenderse a otros dominios.

Los temas ubicados en el cuadrante inferior derecho son considerados temas básicos, ya que poseen alta centralidad, pero baja densidad, lo que indica que son fundamentales para el desarrollo del campo, aunque requieren mayor consolidación interna.

La Inteligencia Artificial se perfila como un área clave en la automatización de la estimación del esfuerzo para el desarrollo de software, pero sus investigaciones aún están dispersas. COCOMO, aunque sigue siendo un referente, muestra señales de obsolescencia frente a modelos más avanzados. Los proyectos de desarrollo de software mantienen su importancia, pero requieren enfoques más específicos alineados con metodologías ágiles y DevOps.

Los modelos de estimación del esfuerzo en el desarrollo de software, tanto tradicionales como modernos, podrían beneficiarse de una mayor



**Fig. 6.** Red de Conexiones Temáticas del Clúster *Estimation*

**Tabla 3.** Número total de documentos del clúster *Estimation*

Nodos	No. de documentos
Estimation	468
Effort Estimation	207
Software Project	143
Cost Estimations	116
Software Development	108
Software Cost Estimations	101
Costs	94
Software Effort Estimation	93
Software Development Effort	63
Estimation Models	56
Software Estimation	53

especialización en subtemas como modelos predictivos e híbridos.

Finalmente, la gestión de proyectos es un punto de conexión crucial entre las técnicas de estimación del esfuerzo en el desarrollo de software y su aplicación práctica, pero la falta de cohesión en la investigación sugiere la necesidad de enfoques más integrados.

Finalmente, los temas ubicados en el cuadrante inferior izquierdo representan áreas poco

desarrolladas y de baja relevancia estratégica, debido a su baja centralidad y baja densidad. *Software Measurement, Information Systems, Requirement Engineering, Project Cost Estimation* y *COCOMO* tienen un desarrollo limitado y una escasa integración dentro de la red temática.

Para cada clúster identificado o clasificado, es posible generar redes de clústeres (*Cluster Networks*). Cada una de estas redes representa, a través de un diagrama, las conexiones e interrelaciones entre el tema central y otros temas relacionados dentro del campo de estudio.

La Fig. 6 muestra el tema central *Estimation* actuando como el núcleo que integra otros conceptos relacionados en el análisis temático. El tamaño de cada nodo indica la importancia relativa de cada subtema dentro del clúster. Subtemas como *Software Development, Cost Estimations* y *Effort Estimation* destacan por su mayor relevancia o volumen de publicaciones relacionadas. Por otro lado, subtemas como *Estimation Models* y *Data Sets*, aunque menos destacados, muestran conexiones específicas con el tema central.

Las líneas que conectan los nodos representan las relaciones entre los diferentes temas y el grosor de estas representa la intensidad de la conexión. La mayor cantidad de líneas conectadas al nodo *Estimation* indica que está fuertemente interconectado con los subtemas *Software, Cost Estimations, Estimation Models, Effort Estimation*. Esto sugiere que los estudios en estimación del esfuerzo en el desarrollo de software tienen una fuerte orientación hacia la estimación de costos. Lo que podría indicar que la mayoría de los enfoques y modelos de predicción utilizados en este ámbito se centran en medir y controlar los costos dentro de los proyectos de software.

El nodo *Software Development* destaca por su tamaño y múltiples conexiones, lo que indica una estrecha relación entre las prácticas y metodologías de desarrollo de software y los procesos de estimación del esfuerzo. Por su parte, los subtemas *Effort Estimation* y *Cost Estimations* presentan conexiones significativas con el nodo central y otros subtemas.

Refuerzan la idea de que la estimación del esfuerzo en el desarrollo de software no solo se basa en cálculos generales, sino que depende de métricas clave como el esfuerzo requerido y los costos involucrados.

**Tabla 4.** Relación entre centralidad y densidad en los clústeres analizados del subperíodo 2014 - 2019

Clúster	Centralidad	Densidad
Effort Estimation	42.90	4.60
Neural Networks	41.03	17.71
Software Reliability	20.18	13.66
Optimization	16.61	8.32
Requirement Engineering	15.07	8.32
Estimation	13.37	3.81
Function Point	12.79	1.53
Software Cost Estimations	12.36	7.89
Software Development Life Cycle	12.32	1.59
Use Cases	10.03	19.14
Artificial Intelligence	8.04	4.18
Software Development Effort Estimation	6.99	13.17
Mobile Applications	3.44	5.6
User Stories	2.09	6.32
Capture Recapture	1.73	56.25
Value Based Software Engineering	1.22	26.24
Embedded Systems	0.66	5.96
Predictive Models	0.65	7.88

Finalmente, subtemas como *Data Sets* y *Estimation Models*, aunque más pequeños, representan áreas metodológicas y prácticas que sustentan el desarrollo del tema central. La inclusión de *Data Sets* sugiere que las estimaciones pueden depender en gran medida de datos previos o históricos.

La Tabla 3 presenta información sobre los nodos del clúster, específicamente el número de documentos asociados a cada nodo. Cada nodo representa un subtema relacionado con *Estimation*, y su importancia relativa dentro del clúster se refleja en la cantidad de documentos.

Siendo el nodo *Estimation* el tema principal del clúster, con la mayor cantidad de documentos asociados (468 documentos), indica que es el tema más investigado y el núcleo alrededor del cual se agrupan los demás subtemas. Su alto volumen de documentos refuerza su posición estratégica como el eje central del campo de investigación.

Los subtemas con un alto número de documentos reflejan áreas de investigación ampliamente estudiadas dentro del campo de estimación del esfuerzo en el desarrollo de

software; *Effort Estimation* (207 documentos) es el segundo nodo más importante, reflejando un enfoque significativo en los estudios relacionados con la medición y predicción del esfuerzo necesario en proyectos de software. *Software Project* (143 documentos) representa investigaciones centradas en la estimación aplicada al desarrollo y gestión de proyectos de software y *Cost Estimations* (116 documentos) enfatiza la importancia de las técnicas y metodologías para calcular costos en el desarrollo de software.

Por otro lado, los subtemas con una menor cantidad de documentos indican áreas con menor volumen de publicaciones, lo que podría significar que son temas emergentes, especializados o menos explorados dentro del contexto de la estimación del esfuerzo en el desarrollo de software. *Software Development* (108 documentos) relaciona las prácticas de desarrollo con los procesos de estimación, *Software Cost Estimations* (101 documentos) sugiere un enfoque en herramientas o software diseñados específicamente para la estimación de costos. El nodo *Costs* (94 documentos) aborda específicamente los conceptos de costos. Finalmente, *Software Effort Estimation* (93 documentos) se centra en herramientas específicas para la estimación del esfuerzo en proyectos para el desarrollo de software.

#### 4.2.2. Análisis del subperíodo 2014-2019

La Tabla 4 revela cambios significativos en la estructura temática del campo de estimación en ingeniería de software en este subperíodo. Se observa la persistencia de algunos temas fundamentales, la aparición de nuevas líneas de investigación y la desaparición de enfoques tradicionales. Estos cambios reflejan la evolución de la disciplina y la consolidación de nuevos paradigmas en la estimación de software.

El clúster *Estimation* continúa como un tema central, aunque su centralidad disminuye, lo que sugiere que ha pasado de ser un eje estratégico a un área menos interconectada dentro de la red temática. Los *Function Point*, ampliamente utilizados en metodologías de estimación del esfuerzo, mantienen su presencia, pero con una reducción en su centralidad y una disminución en su densidad, lo que sugiere una menor cohesión

interna en su desarrollo. Algunos clústeres se mantienen en los subperíodos 2008-2013 y 2014-2019, lo que indica que siguen siendo relevantes en la investigación del área.

El clúster *Software Reliability* sigue siendo un tema recurrente, con una densidad alta y centralidad estable, lo que indica que continúa desempeñando un papel relevante en la investigación de calidad del software. Sin embargo, la *Artificial Intelligence* experimenta una disminución en su centralidad, posiblemente debido a la fragmentación de la investigación en esta área, derivando en estudios más específicos. Por otro lado, los *Embedded Systems* se mantienen presentes, pero con una reducción en su centralidad y densidad, lo que refleja una disminución en su importancia dentro del campo de estimación del esfuerzo en el desarrollo de software. Finalmente, *Software Cost Estimations* presenta una variación en su densidad, lo que sugiere una maduración progresiva del tema.

En contraste, varios clústeres emergen en el subperíodo 2014-2019, lo que indica la aparición de nuevas áreas de interés en la estimación del esfuerzo en el desarrollo de software. *Effort Estimation*, que no estaba presente en el primer subperíodo, se convierte en el clúster central y estratégico en el segundo subperíodo, lo que refleja su creciente relevancia en la industria del software. De manera similar, *Neural Networks* surge como un tema clave, lo que indica un interés creciente en la aplicación de inteligencia artificial a los procesos de estimación del esfuerzo.

Además, el tema *Optimization* aparece como un nuevo clúster, lo que sugiere una mayor exploración de técnicas de optimización en la estimación del esfuerzo. *Use Cases* se introduce como un tema bien cohesionado, lo que indica su consolidación en la investigación. El *Software Development Life Cycle* también aparece, lo que sugiere un interés en la estimación del esfuerzo aplicada a modelos de desarrollo colaborativo. La inclusión de *Mobile Applications* refleja la importancia creciente de la estimación del esfuerzo en entornos móviles, mientras que el tema de *User Stories* indica una mayor adopción de metodologías ágiles en los procesos de estimación del esfuerzo.

Asimismo, el tema *Capture Recapture* aparece con un enfoque en la calidad del software, evidenciando el interés en la relación entre estimación del esfuerzo y control de calidad. Finalmente, *Value Based Software Engineering* surge con una alta cohesión interna, lo que indica un desarrollo estructurado en la evaluación del valor agregado en los procesos de estimación del esfuerzo.

Paralelamente, varios clústeres que estaban presentes en el período 2008-2013 desaparecen en 2014-2019, lo que indica que han perdido relevancia o han sido absorbidos por otras líneas de investigación. Un caso notable es la desaparición de *Softcomputing Techniques*, lo que sugiere que este enfoque ha sido reemplazado o absorbido por metodologías más especializadas como *Neural Networks* y *Predictive Models*.

De manera similar, clústeres como la *Project Managers*, *predicción* y *COCOMO* dejan de ser representativos, lo que sugiere una transición desde modelos generales de estimación del esfuerzo hacia enfoques más específicos, como *Effort Estimation* y *Optimization*. Otros temas, como *Software Measurement* y *Functional Size*, también desaparecen, lo que indica una menor dependencia de estas métricas tradicionales dentro de los modelos de estimación del esfuerzo contemporáneos. Asimismo, los clústeres de *Requirement Engineering* y *Nonlinear Systems* dejan de figurar, lo que sugiere que han sido absorbidos en áreas más amplias y especializadas.

Se observa, por lo tanto, una transición desde enfoques tradicionales hacia metodologías más especializadas. La consolidación de la *Effort Estimation* como el tema más importante en el segundo período evidencia un cambio en la forma en que la estimación del esfuerzo es abordada en la industria del software. Además, la desaparición de *Softcomputing Techniques* y la aparición de *Neural Networks* y *Predictive Models* reflejan la creciente integración de inteligencia artificial en la estimación del esfuerzo.

Otro cambio relevante es la inclusión de *User Stories* y *Use Cases*, lo que sugiere que la estimación del esfuerzo en metodologías ágiles ha adquirido una mayor relevancia. Finalmente, la ausencia de *COCOMO* y *Software Measurement* señala un declive en el uso de estos modelos

**Tabla 5.** Propiedades del clúster *estimación del esfuerzo*

Propiedad	Valor
Número total de documentos	193
h-index	25
Total de citas	2,792

tradicionales, en favor de enfoques más dinámicos y adaptativos.

Los valores presentados en la Tabla 5 reflejan que la estimación del esfuerzo es un área de estudio madura y consolidada. Su alta cantidad de documentos publicados, combinada con un índice h elevado y un gran volumen de citas, sugiere que este clúster no solo ha sido ampliamente investigado, sino que también ha generado contribuciones clave para la evolución del campo.

Estos resultados indican que el estudio de la estimación del esfuerzo sigue siendo un tema relevante, con un impacto significativo en la literatura. Su crecimiento y desarrollo sugieren que sigue siendo un campo de oportunidad para nuevas investigaciones, particularmente en la aplicación de técnicas avanzadas como inteligencia artificial, optimización y modelos predictivos en los procesos de estimación del esfuerzo.

Además, la alta citación de los trabajos dentro del clúster implica que existe una base teórica y metodológica sólida, lo que proporciona un punto de partida robusto para futuras investigaciones. La evolución de este campo podría estar orientada hacia la integración de enfoques más modernos, como el aprendizaje automático y la automatización de procesos en entornos ágiles y DevOps, lo que fortalecería aún más su relevancia dentro de la disciplina.

La Fig. 7 presenta el mapa estratégico del clúster *Effort Estimation*. El clúster *Effort Estimation* es el tema central del análisis en este subperíodo, con la mayor centralidad e impacto bibliométrico (h-index = 25), lo que indica su consolidación como un eje fundamental en la investigación de estimación del esfuerzo en ingeniería de software. A su vez, *Software Reliability* (h-index = 19) y *Optimization* (h-index = 17) también se posicionan como áreas clave, aunque su impacto bibliométrico es menor que el

de *Effort Estimation*, su presencia sigue siendo significativa, indicando la importancia de evaluar la calidad del software y mejorar los modelos de estimación del esfuerzo mediante técnicas de optimización. Por otro lado, {Neural Networks} aparece con alta centralidad y densidad, sugiriendo un enfoque emergente en la integración de modelos distribuidos dentro del proceso de estimación del esfuerzo.

Los clústeres *Requirement Engineering* (h-Index = 11), *Software Development Life Cycle* (h-index = 14) y *Estimation* (h-index = 10) son temas estratégicos con alta centralidad, pero con un desarrollo interno moderado, lo que sugiere la necesidad de mayor consolidación teórica y metodológica. *Function Point* (h-index = 7), a pesar de ser una métrica reconocida en la estimación del esfuerzo en el desarrollo del software, muestra baja densidad, lo que indica que su aplicación sigue fragmentada dentro del campo.

Los temas altamente desarrollados pero periféricos incluyen *Use Cases* (h-index = 5), *Software Development Effort Estimation* (h-index = 9) y *Value Based Software Engineering* (h-index = 4), que, aunque presentan cohesión interna, tienen una baja conexión con otros clústeres, lo que indica que son líneas de investigación bien establecidas. A pesar de no tener el h-index más alto en comparación con los otros temas, su influencia estructural en la red temática indica que su relevancia está en crecimiento y podría tener un mayor impacto en investigaciones futuras.

Finalmente, los clústeres emergentes o marginales incluyen *Predictive Models* (h-index = 3), *Embedded Systems* (h-index = 5), *Mobile Applications* (h-index = 4) y *User Stories* (h-Index = 6), que aún tienen un bajo desarrollo y conexión con la red temática, pero representan áreas con potencial de crecimiento, especialmente en el contexto de metodologías ágiles y nuevos entornos de software. En general, estos resultados sugieren una evolución en la estimación del esfuerzo, con un crecimiento de enfoques basados en optimización, inteligencia artificial y software de código abierto, mientras que algunos modelos tradicionales requieren mayor integración para mantenerse relevantes.

La Figura 8 representa la red de clústeres del tema *Effort Estimation*, mostrando las

**Tabla 6.** Número de documentos del clúster *Estimation*

Nodos		No. de documentos
Effort Estimation		265
Lifecycle		88
Software Management	Project	71
Agile Development	Software	69
Use Case Points		56
Software Management	Project	52
Software Development		42
Software Management	Project	42
Effort Estimation Model		35
Estimations Techniques		32
Clustering		12

**Tabla 7.** Relación entre centralidad y densidad en los clústeres analizados del subperiodo 2020 - 2024

Clúster		Centralidad	Densidad
Software Estimation	Effort	57.06	13.05
Decision Trees		28.31	5.55
Software Estimations	Cost	25.15	9.07
Software Metrics		16.65	3.32
Agile Development	Software	15.34	4.33
Software Reliability		13.61	9.62
Features Selection		8.15	8.25
Human Management	Resource	4.19	3.23
Software Metrics		1.67	5.78
Function Point		1.08	6.17
Human Estimations	Posee	0.74	5.69

interconexiones entre sus subtemas clave. Se observa que *Effort Estimation* es el nodo central y más influyente, con múltiples conexiones a temas fundamentales en la gestión y desarrollo de software.

Entre los nodos más relevantes destacan *Software Development*, *Software Development Projects* y *Software Project Management*, lo que evidencia la relación directa entre la estimación del esfuerzo y la planificación de proyectos. La presencia de *Estimations Techniques* y *Effort Estimation Model* sugiere un enfoque metodológico que respalda la toma de decisiones basada en métricas. Además, la conexión con

*Agile Software Development* indica la creciente adopción de metodologías ágiles en los procesos de estimación del esfuerzo. La inclusión de *Lifecycle* sugiere que este campo está en constante evolución, evaluando enfoques previos para mejorar la precisión de la estimación del esfuerzo. Finalmente, la presencia de *Clustering* y *Use Case Points* destaca la integración de técnicas analíticas y basadas en casos prácticos para mejorar la predicción del esfuerzo requerido en proyectos de software. Esta red evidencia que la estimación del esfuerzo es un área multidimensional, con aplicaciones que abarcan desde la gestión de proyectos hasta la optimización de metodologías en el desarrollo de software, consolidándose como un componente esencial en la ingeniería de software moderna.

La Tabla 6 confirma que la estimación del esfuerzo es el núcleo central dentro del clúster, con el mayor número de documentos, reflejando su importancia en la planificación y gestión de proyectos de software.

La fuerte presencia de estudios sobre ciclo de vida, metodologías ágiles y gestión de proyectos sugiere que la estimación del esfuerzo se ha adaptado a enfoques modernos, integrándose con prácticas ágiles y modelos de desarrollo iterativo. Además, la inclusión de nodos como modelos y técnicas de estimación, aunque con menor cantidad de publicaciones, indica un interés en la mejora metodológica y en la aplicación de enfoques analíticos avanzados para optimizar la predicción del esfuerzo requerido. Finalmente, la presencia de clustering como nodo con menor volumen de documentos sugiere que se están explorando nuevas estrategias de clasificación y análisis de datos para mejorar la precisión de las estimaciones. En conjunto, estos resultados reflejan una evolución en el campo, donde la estimación del esfuerzo no solo es un elemento teórico, sino una herramienta clave en la planificación eficiente y el éxito de los proyectos de software.

#### 4.2.3. Análisis del periodo 2020-2024

La Tabla 7 evidencia una evolución significativa en los enfoques y prioridades dentro de la investigación sobre estimación del esfuerzo en desarrollo de software, reflejando un cambio

**Tabla 8.** Propiedades del clúster *Software Effort Estimation*

Propiedad	Valor
Número total de documentos	206
h-index	14
Total de citas	903

progresivo desde metodologías tradicionales hacia enfoques más especializados y automatizados.

Se observa una transición de la estimación del esfuerzo genérica a enfoques específicos. En el primer periodo (2008-2013), el clúster más influyente era *Estimation*, lo que indica que la estimación del esfuerzo era un concepto amplio y dominante. En el segundo periodo (2014-2019), la investigación comenzó a fragmentarse en temas más especializados, donde *Effort Estimation* seguía siendo clave, pero surgieron enfoques complementarios como *Neural Networks*, *Optimization* y *Software Reliability*. En el tercer periodo (2020-2024), se observa un refinamiento aún mayor, con un enfoque en herramientas automatizadas, destacando *software para la estimación del esfuerzo* como el clúster con mayor centralidad, lo que sugiere un creciente protagonismo de la automatización y las herramientas basadas en software en los procesos de estimación del esfuerzo.

Además, se evidencia un desplazamiento de los modelos tradicionales hacia técnicas avanzadas. En el primer periodo, la estimación del esfuerzo estaba dominada por modelos clásicos como *COCOMO* y *Function Point*. En el segundo periodo, estos enfoques comenzaron a perder relevancia en favor de técnicas más avanzadas, como *Neural Networks* y *Optimization*. En el tercer periodo, se consolida el uso de inteligencia artificial, con *Decision Trees* como uno de los clústeres más relevantes, lo que indica un desplazamiento hacia modelos predictivos basados en aprendizaje automático.

A lo largo de los tres periodos, *Software Reliability* ha mantenido una presencia constante, aunque con una leve disminución en su centralidad. Sin embargo, su alta densidad indica que sigue siendo un campo de investigación

maduro y relevante para mejorar la precisión de los modelos de estimación. Asimismo, han emergido temas como *Features Selection* y *Parameters Estimation*, lo que refleja un interés creciente en la mejora de la precisión de las estimaciones mediante preprocesamiento de datos y optimización de modelos.

Algunos temas han mostrado persistencia, mientras que otros han desaparecido. *Agile Software Development* ha mantenido su presencia a lo largo de los tres periodos, reflejando su adopción sostenida en la industria. En contraste, temas como *Requirements Engineering* han desaparecido en el periodo 2020-2024, lo que sugiere que su impacto en la estimación del esfuerzo ha sido absorbido por enfoques más automatizados.

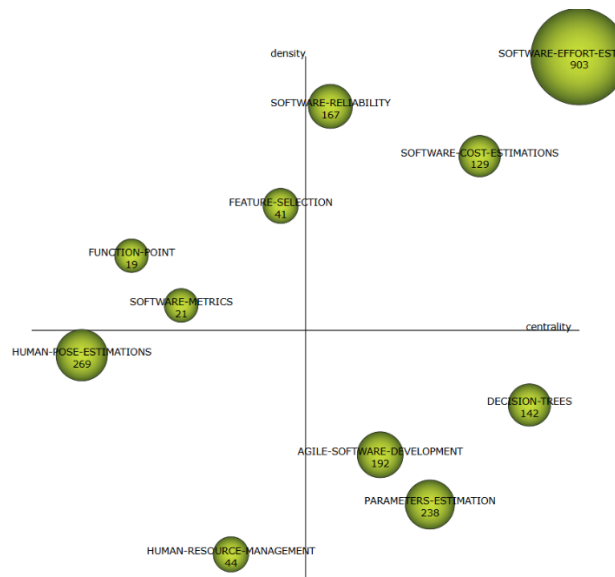
Estos hallazgos reflejan una clara evolución en la estimación del esfuerzo, con una transición hacia enfoques híbridos que combinan modelos tradicionales con inteligencia artificial y aprendizaje automático, optimizando así la precisión y eficiencia de la planificación en proyectos de software.

La Tabla 8 presenta las propiedades bibliométricas del clúster *Software Effort Estimation*, evidenciando su relevancia dentro de la investigación en estimación del esfuerzo en el desarrollo de software.

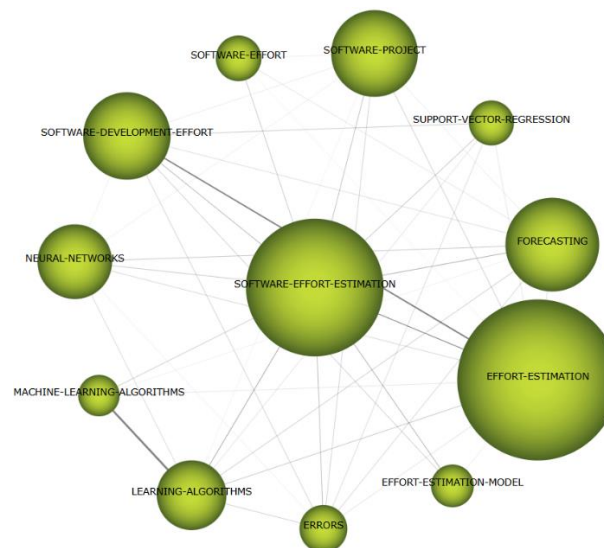
El número total de documentos analizados (206) indica una producción científica considerable en esta área, lo que sugiere un creciente interés por las herramientas y plataformas automatizadas para la estimación del esfuerzo en proyectos de software. La presencia de un indicador de impacto científico (h-index) de 14 muestra que al menos 14 documentos dentro del clúster han sido citados por lo menos 14 veces, lo que denota una influencia moderada dentro de la literatura académica.

Además, el total de citas recibidas (903) indica que los estudios relacionados con software para la estimación del esfuerzo han generado un impacto significativo, reflejando su importancia en la comunidad científica. La cantidad de citas sugiere que los modelos y herramientas de software en este ámbito están siendo ampliamente referenciados, consolidando su papel como un área clave en la evolución de la estimación del esfuerzo en ingeniería de software.





**Fig. 9.** Mapa estratégico del clúster *Software Effort Estimation* basado en centralidad, densidad y número total de citas



**Fig. 10.** Red de Conexiones Temáticas del Clúster *Software Effort Estimation*

Esta tendencia sugiere un desplazamiento desde modelos tradicionales hacia enfoques basados en software y automatización, con un impacto creciente en la planificación y optimización de recursos en el desarrollo de software.

La figura 9 presenta el mapa estratégico del clúster *software para la estimación del esfuerzo* el cual refleja la distribución de los temas en función de su centralidad, densidad y volumen de documentos analizados, lo que permite identificar su evolución y relevancia dentro del campo.

En el cuadrante superior derecho, se encuentra *Software Effort Estimation*, que presenta la mayor centralidad y densidad, consolidándose como un tema motor clave en la investigación. Junto a este, destacan *Software Cost Estimations* y *Decision Trees*, lo que sugiere que la integración de modelos predictivos y herramientas de estimación automatizadas juega un papel central en la planificación y optimización de recursos en proyectos de software.

En el cuadrante superior izquierdo, se ubican temas altamente desarrollados, pero con menor conexión con la red global, como *Software Reliability* y *Features Selection*, indicando que estos aspectos han sido ampliamente estudiados internamente, pero aún pueden vincularse más con otros campos. En el cuadrante inferior derecho, emergen tendencias en crecimiento como *Agile Software Development* y *Parameters Estimation*, lo que sugiere que la investigación en metodologías ágiles y la optimización de modelos de estimación continúa expansionándose. Finalmente, en el cuadrante inferior izquierdo, se encuentran temas con alta densidad, pero baja centralidad, como *Human Posee Estimations* y *Human Resource Management*, que, aunque son áreas bien desarrolladas, presentan menor influencia en la red central del clúster.

Estos resultados reflejan una clara evolución en la estimación del esfuerzo, con una transición hacia enfoques híbridos que combinan herramientas automatizadas, modelos predictivos y metodologías ágiles para mejorar la precisión en la planificación y ejecución de proyectos de software.

La Fig. 10 muestra la red de clústeres (Cluster's Network) del tema central *Software Effort Estimation*, representando sus relaciones con otros temas relacionados dentro del campo de estudio.

En comparación con los subperíodos anteriores, se observa la evolución de las estimaciones en el esfuerzo de software. Para el periodo 2008-2013 predominan enfoques tradicionales como *Software Effort Estimation*, *Software Development Effort* y *Forecasting*, reflejando el uso de modelos deterministas como *COCOMO* y *Funtion Points*. En el periodo 2014-2019, se reflejó la consolidación de técnicas avanzadas como *Neural Networks*, *Machine*

**Table 9.** Red de Conexiones Temáticas del Clúster *Software Effort Estimation*

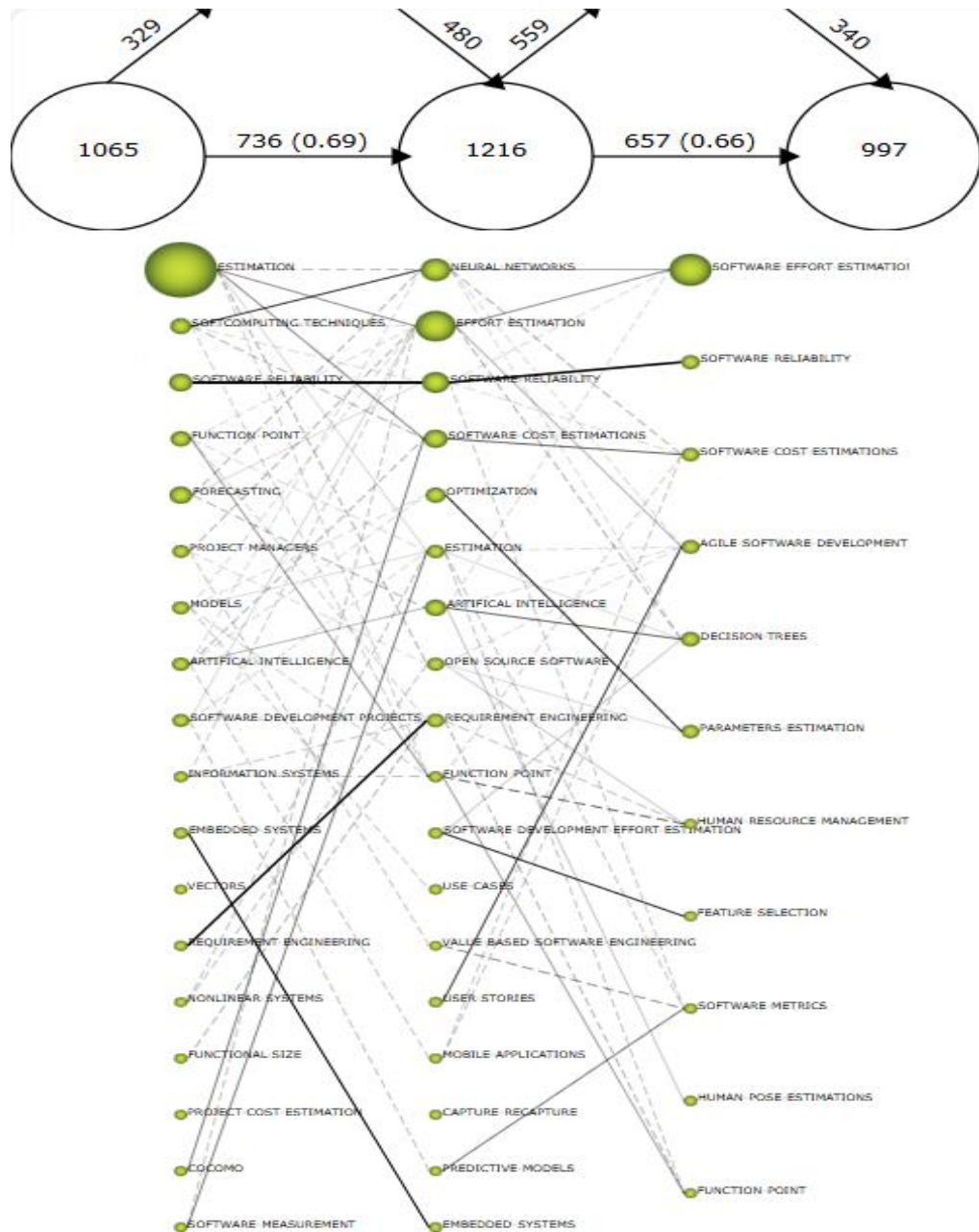
Nodos	No. de documentos
Effort Estimation	182
Software Effort Estimation	149
Forecasting	87
Software Development Effort	78
Software Project	77
Neural Networks	60
Learning Algorithms	53
Errors	22
Software Effort	19
Support Vector Regression	18
Effort Estimation Model	15
Machine Learning Algorithms	13

*Learning Algorithms* y *Support Vector Regression*, indicando la adopción de enfoques de inteligencia artificial en la estimación del esfuerzo. En el periodo 2020-2024 conceptos como *Effort Estimations Models* y *Errors* sugieren un interés creciente en la optimización de modelos predictivos, mientras que la fuerte conexión entre *Software Effort Estimation* y *Effort Estimation* destacan la centralidad de la estimación del esfuerzo como tema clave en la evolución del campo.

La Tabla 9 presenta información detallada sobre los nodos del clúster *Software Effort Estimation* específicamente el número de documentos asociados a cada nodo. Cada nodo representa un subtema relacionado con *Estimation*, y su importancia relativa dentro del clúster se refleja en la cantidad de documentos.

### 4.3. Evolución del campo de estudio

La Fig. 11 ilustra la evolución temática del campo de investigación de las estimaciones (2008-2024) mediante un análisis de fracciones de solapamiento entre subperíodos sucesivos, reflejando la continuidad y transformación de los temas de investigación a lo largo del tiempo.



**Fig. 11.** Evolución temática de las estimaciones (2008-2024) y el análisis de fracciones de solapamiento entre subperiodos

En la parte superior de la figura, los círculos representan los tres subperiodos analizados, con el número total de documentos en cada fase. Los valores en paréntesis indican el grado de

solapamiento entre subperiodos consecutivos, evidenciando qué porcentaje de los temas previos se mantiene en el siguiente periodo. Se observa una transición progresiva con una tasa de

solapamiento relativamente constante (0.69 y 0.66), lo que sugiere una evolución estable del campo de investigación sin cambios disruptivos.

En la parte inferior de la figura, se representa la conexión de palabras clave entre los distintos subperíodos, las líneas gruesas que indican una fuerte continuidad temática y líneas más delgadas que representan conexiones emergentes o en declive.

Se identifican temas consolidados como *Estimation*, *Neural Networks*, *Optimization* y *Software Cost Estimations*, que han mantenido su presencia en los distintos periodos. Asimismo, nuevas tendencias como *Decision Trees*, *Parameters Estimation*, *Features Selection* y *Agile Software Development* han ido adquiriendo relevancia en los periodos más recientes, indicando un cambio hacia modelos híbridos que combinan inteligencia artificial y metodologías ágiles.

Finalmente, la figura también sugiere una reducción en la centralidad de ciertos temas tradicionales, como *Function Point* y *Embedded Systems*, que han perdido protagonismo en los últimos periodos. Esto confirma una transición del uso de modelos convencionales hacia enfoques basados en técnicas automatizadas y de aprendizaje automático para la estimación del esfuerzo en ingeniería de software.

## 5. Conclusiones y trabajos futuros

Este estudio bibliométrico ha permitido analizar la evolución de la estimación del esfuerzo en el desarrollo de software, identificando tendencias clave, clústeres temáticos y redes de colaboración científica. A lo largo de los periodos analizados, se observa una transición desde métodos tradicionales como *COCOMO* y *Function Point* hacia enfoques híbridos basados en *Artificial Intelligence*, *Machine Learning Algorithms* y *Optimization*. La consolidación del clúster *Software Effort Estimation* como el más influyente en el periodo 2020-2024 sugiere un enfoque creciente en herramientas automatizadas y modelos predictivos. Asimismo, el análisis de fracciones de solapamiento muestra una evolución temática estable, con la integración progresiva de nuevas áreas de investigación. En términos de

colaboración científica, se identificaron autores e instituciones clave que han liderado el campo, destacando la contribución de investigadores como Mendes, Boehm y Jørgensen, así como la influencia de países como Estados Unidos, China y Alemania. La visualización de redes de coautoría y citas evidenció una comunidad académica activa, con tendencias emergentes en metodologías ágiles, DevOps y optimización basada en datos.

Para futuras investigaciones, se recomienda profundizar en la integración de modelos híbridos que combinen enfoques tradicionales y técnicas avanzadas de aprendizaje automático, con el objetivo de mejorar la precisión y eficiencia en la estimación del esfuerzo en proyectos de software. Además, es necesario explorar el impacto de la automatización en entornos de desarrollo ágil y DevOps, así como evaluar la aplicabilidad de estos modelos en diferentes contextos organizacionales. Un área clave de oportunidad es el desarrollo de herramientas de estimación adaptativas, que incorporen técnicas de optimización y análisis de datos en tiempo real para mejorar la toma de decisiones en la gestión de proyectos. Finalmente, se sugiere ampliar el análisis a nivel regional para identificar patrones de adopción de metodologías de estimación del esfuerzo en distintas industrias y contextos, lo que permitirá generar estrategias más efectivas para la implementación de estas técnicas en el desarrollo de software a nivel global.

## Agradecimientos

Esta investigación fue apoyada por la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) con la beca No. CVU: 328123. Los autores reconocen al Tecnológico Nacional de México y al Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET) el apoyo prestado para la realización del presente trabajo.

## Referencias

1. Abulalqader, F., Ali, A. (2018). Comparing different estimation methods for software effort. 2018 1st Annual International

- Conference on Information and Sciences (AiCIS), IEEE, pp. 13–22.
2. **Albrecht, A. J., Gaffney, J. E. (1983).** Software function, source lines of code, and development effort prediction: A software science validation. *IEEE Transactions on Software Engineering*, Vol. 9, No. 6, pp. 639–648.
  3. **Alvarez-Marin, A., Castillo-Vergara, M., Geldes-González, C. (2017).** Análisis bibliométrico de la realidad aumentada y su relación con la administración de negocios. *Información Tecnológica*, Vol. 28, No. 4, pp. 57–66.
  4. **Aurora, M., Mallick, P. K., Pradhan, C., Singh, P., Dash, R. (2020).** A systematic literature review of machine learning estimation approaches in scrum projects. In **Mallick, P. K., Behera, C., Balusamy, B., Nayak, R.**, editors, *Cognitive Informatics and Soft Computing*. Springer, Singapore, pp. 573–586.
  5. **Baykoucheva, S. (2019).** Eugene Garfield's ideas and legacy and their impact on the culture of research. *Publications*, Vol. 7, No. 2, pp. 1–12.
  6. **Bhirud, S. G., Deshpande, M. V. (2010).** Analysis of combining software estimation techniques. *International Journal of Computer Applications*, Vol. 5, No. 3, pp. 1–2.
  7. **Cooper, R. E. (1992).** The complete Cocomo model: Basic, intermediate, detailed, and incremental versions for the original, enhanced, Ada, and Ada process models of Cocomo. In **Gulledge, T. R., Hutzler, W. P., Lovelace, J. S.**, editors, *Cost estimating and analysis*. Springer, pp. 111–149.
  8. **Dalkey, N., Helmer, O. (1963).** An experimental application of the delphi method to the use of experts. *Management Science*, Vol. 9, No. 3, pp. 458–467.
  9. **Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., Lim, W. M. (2021).** How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, Vol. 133, pp. 285–296.
  10. **Fitzgerald, B. (1997).** Systems development methodologies. In **Wojtkowski, W. G. et al.**, editors, *Proceedings of the International Conference on Information Systems Development*. Springer US, pp. 127–139.
  11. **Gamboa-Cruzado, J., Cuya-Chuica, L., López-Goycochea, J., Nuñez-Meza, A., Del-Valle-Jurado, C. (2024).** Impact of 5g technology on cybersecurity: A comprehensive systematic and bibliometric review. *Computación y Sistemas*, Vol. 28, No. 2, pp. 367–386.
  12. **Garfield, E. (1955).** Citation indexes for science. *Science*, Vol. 122, No. 3159, pp. 108–111.
  13. **Haugen, N. C. (2006).** An empirical study of using planning poker for user story estimation. *AGILE 2006 (AGILE'06) Proceedings, IEEE*, pp. 9–34.
  14. **Jorgensen, M., Molokken-Ostvoid, K. (2004).** Reasons for software effort estimation error: impact of respondent role, information collection approach, and data analysis method. *IEEE Transactions on Software Engineering*, Vol. 30, No. 12, pp. 993–1007.
  15. **Jørgensen, M., Sjøberg, D. I. K. (2001).** Impact of effort estimates on software project work. *Information and Software Technology*, Vol. 43, No. 15, pp. 939–948.
  16. **Kokol, P., Zagoranski, S., Kokol, M. (2020).** Software development with scrum: A bibliometric analysis and profile. *Library Philosophy and Practice (e-journal)*, No. 4705.
  17. **Moløkken, K., Jørgensen, M. (2003).** A review of surveys on software effort estimation. *Proceedings of the 2003 International Symposium on Empirical Software Engineering (ISESE 2003)*, IEEE, pp. 223–230.
  18. **Olvera, C., Berbegal-Mirabent, J., Merigó, J. M. (2018).** A bibliometric overview of university business collaboration between 1980 and 2016. *Computación y Sistemas*, Vol. 22, No. 4, pp. 1171–1190.
  19. **Rajlich, V. (2014).** Software evolution and maintenance. *Future of Software Engineering Proceedings (FOSE 2014)*, Association for

Computing Machinery, New York, NY, USA, pp. 133–144.

20. **Sharma, B., Purohit, R. (2018).** Review of current software estimation techniques. In **Panda, B., Sharma, S., Roy, N. R.**, editors, *Data Science and Analytics*. Springer, Singapore, pp. 380–399.
21. **Sousa, A. O., dos Santos, R. P., Lima, G. A., Oliveira, A. S. (2023).** Applying machine learning to estimate the effort and duration of individual tasks in software projects. *IEEE Access*, Vol. 11, pp. 89933–89946.
22. **Spinak, E. (2017).** In memoriam: Eugene Garfield (1925–2017). *SciELO en Perspectiva*. Recuperado el 7 de noviembre de 2024, de <https://blog.scielo.org/es/2017/03/03/inmemoriam-eugene-garfield-1925-2017/>.
23. **Tawosi, V., Moussa, R., Sarro, F. (2022).** Agile effort estimation: Have we solved the problem yet? insights from a replication study. *arXiv*.
24. **Teslyuk, V., Batyuk, A., Voityshyn, V. (2022).** Method of software development project duration estimation for scrum teams with differentiated specializations. *Systems*, Vol. 10, No. 4, pp. 123.

*Article received on 08/02/2025; accepted on 30/05/2025.  
\*Corresponding author is Noé Alejandro Castro Sánchez.*