

2017



DOCUMENTO FINAL DE LA PASANTÍA DE INVESTIGACIÓN

**PERFIL DE LA INVESTIGACIÓN CIENTÍFICA EN
EL ÁREA DE CONOCIMIENTO DE LA
INGENIERÍA DE SOFTWARE DENOMINADA
“CALIDAD DEL SOFTWARE”**

SANTIAGO DE JESUS MARTINEZ RODRIGUEZ

**UNIVERSIDAD DEL MAGDALENA
FACULTAD DE INGENIERÍA
PROGRAMA DE INGENIERÍA DE SISTEMAS
SANTA MARTA D.T.C.H.
2017**



**PERFIL DE LA INVESTIGACIÓN CIENTÍFICA EN EL ÁREA DE CONOCIMIENTO DE LA
INGENIERÍA DE SOFTWARE DENOMINADA “CALIDAD DEL SOFTWARE”**

SANTIAGO DE JESUS MARTINEZ RODRIGUEZ

Trabajo presentado para optar por el título de Ingeniero de Sistemas

**DIRECTOR
ERNESTO AMARU GALVIS LISTA MSC. PHD.**

**UNIVERSIDAD DEL MAGDALENA
FACULTAD DE INGENIERÍA
PROGRAMA DE INGENIERIA DE SISTEMAS
SANTA MARTA D.T.C.H
2017**



Tabla de contenido

1. Índice de ilustraciones.....	3
2. Índice de tablas.....	4
3. Introducción	4
4. Marco teórico	5
4.1 Ingeniería del software	6
4.2 Calidad del software	7
4.3 Bibliometría	7
5. Metodología.....	7
5.1 Inteligencia	8
5.2 Análisis y diseño	9
5.3 Selección	10
6. Resultados.....	10
6.1 Análisis de primer orden	10
6.1.1 Documentos por año.....	11
6.1.2 Producción de documentos por tipo	12
6.1.3 Top 10 de instituciones con la mayor producción (Scopus)	16
6.1.4 Top 10 de autores con mayor producción de documentos seleccionados para el análisis.....	18
6.1.5 Top 10 de documentos por países(Scopus)	21
6.1.6 Top 10 de los documentos más citados (Scopus)	23
6.1.7 Top 10 de las palabras claves más utilizadas(Scopus).....	25
6.1.8 Top 10 de las palabras más utilizadas de la bolsa de palabras	26
6.2 Análisis de segundo orden.....	30
6.2.1 Top 10 red de concurrencia de autores	30
6.2.2 Top 10 red de concurrencia de afiliaciones(Scopus)	31
6.2.3 Top 10 red de concurrencia de palabras claves(Scopus)	32
6.2.4 Top 10 red de concurrencia de la bolsa de palabras	33
7. Discusión	35
8. Conclusiones	36
9. Agradecimientos	37
10. Bibliografía.....	37
11. Anexos.....	39



1. Índice de ilustraciones

Ilustración 1	Cantidades de documentos por año.....	12
Ilustración 2	Distribución de los documentos en estudio, publicados en Scopus por tipo	13
Ilustración 3	Distribución de los documentos en estudio, publicados en Scopus por tipo y por año.....	14
Ilustración 4	Distribución de los documentos en estudio, publicados en ISI Web por tipo.....	15
Ilustración 5	Distribución de los documentos en estudio, publicados en ISI Web por tipo y por año.....	15
Ilustración 6	Top 10: Top 10 de instituciones con la mayor producción de los documentos en estudio, publicados e Scopus, por año.....	17
Ilustración 7	Distribución de los documentos publicados por el Top 10 de instituciones con la mayor producción de documentos en estudio, publicados en Scopus, por año	17
Ilustración 8	Distribución del Top 10 de autores con mayor producción de documentos seleccionados publicados en Scopus.....	19
Ilustración 9	Distribución del Top 10 de autores con mayor producción de documentos seleccionados publicados en ISI Web.....	20
Ilustración 10	Distribución del Top 10 de autores con mayor producción de documentos seleccionados publicados en Scopus, por año.....	20
Ilustración 11	Distribución del Top 10 de autores con mayor producción de documentos seleccionados publicados en ISI Web, por año	21
Ilustración 12	Distribución del Top 10 de los países con mayor producción de documentos seleccionados publicados en Scopus.....	22
Ilustración 13	Top 10 de los países con mayor producción de documentos seleccionados publicados en Scopus, por año	23
Ilustración 14	Top 10 de los documentos más citados (Scopus)	24
Ilustración 15	Top 10 de las palabras claves más utilizadas (Scopus)	26
Ilustración 16	Top 10 de las palabras claves más utilizadas por año (Scopus).....	26
Ilustración 17	Top 10 de las palabras más utilizadas de la bolsa de palabras (Scopus)	28
Ilustración 18	Top 10 de las palabras más utilizadas de la bolsa de palabras (ISI Web)	28
Ilustración 19	Top 10 de las palabras más utilizadas de la bolsa de palabras en los años (Scopus)	29
Ilustración 20	Top 10 de las palabras más utilizadas de la bolsa de palabras en los años (ISI Web).....	29
Ilustración 21	Top 10 red de concurrencia de autores (Scopus).....	30
Ilustración 22	Top 10 red de concurrencia de autores (ISI Web)	31
Ilustración 23	Top 10 red de concurrencia de afiliaciones (Scopus)	32
Ilustración 24	Top 10 red de concurrencia de palabras claves (Scopus)	33
Ilustración 25	Top 10 red de concurrencia de la bolsa de palabras (Scopus)	34
Ilustración 26	Top 10 red de concurrencia de la bolsa de palabras (ISI Web)	34



2. Índice de tablas

Tabla 1 Áreas de conocimiento (SWEBOK)	6
Tabla 2 Primera ecuación de búsqueda (Elaboración propia)	8
Tabla 3 Ecuación final para ambas bases de datos (Elaboración propia).....	8
Tabla 4 Resumen de resultados.....	10
Tabla 5 Número de documentos por año (Scopus)	11
Tabla 6 Distribución de los documentos en estudio, publicados en Scopus por tipo y por año	12
Tabla 7 Distribución de los documentos en estudio, publicados en ISI Web por tipo y por año.	14
Tabla 8 Top 10 de instituciones con la mayor producción de los documentos en estudio, publicados en Scopus, por año	16
Tabla 9 Top 10 de autores con mayor producción de documentos seleccionados publicados en Scopus, por año.....	18
Tabla 10 Top 10 de autores con mayor producción de documentos seleccionados publicados en ISI Web, por año	19
Tabla 11 Top 10 de los países con mayor producción de documentos seleccionados publicados en Scopus, por año.....	22
Tabla 12 Top 10 de los documentos más citados (Scopus).....	24
Tabla 13 Top 10 de las palabras claves más utilizadas (Scopus).....	25
Tabla 14 Top 10 de las palabras más utilizadas de la bolsa de palabras (Scopus)	27
Tabla 15 Top 10 de las palabras más utilizadas de la bolsa de palabras (ISI Web).....	27

3. Introducción

La Universidad del Magdalena (Unimagdalena) realiza su actividad investigativa en los grupos de investigación como parte del Sistema Nacional de Ciencia, Tecnología e Innovación (SNCTI). Particularmente, se promueve el desarrollo de proyectos de investigación enfocados a dar solución a problemáticas del entorno y el desarrollo humano, social, tecnológico y económico de la región y del país [1]. En esta actividad también se articula la formación de alto nivel ofrecida en los programas de maestría y doctorado.

La orientación de los proyectos de investigación en Unimagdalena también se prioriza teniendo en cuenta las tendencias en la producción académica en ciertas áreas de interés y en las que se tienen capacidades de investigación. Específicamente, dentro del proyecto de investigación estratégico con el cual se está construyendo el Plan Prospectivo para el desarrollo de Ciencia, Tecnología e Innovación de la Universidad del Magdalena [2], se tiene prevista una fase en la cual se identificarán las tendencias de la investigación en un conjunto de áreas de conocimiento que son estratégicas para la institución. En esta fase, se realizarán ejercicios de vigilancia científica y tecnológica en los cuales se analizarán metadatos de las publicaciones científicas indexadas en base de datos como “SCOPUS” [3] e “ISI Web of knowledge” [4], para identificar áreas y temas de investigación importantes en los que se deberían enfocar los investigadores y grupos de la universidad.

La realización de este tipo de ejercicios de vigilancia científica y tecnológica requiere de ciertas capacidades tecnológicas, como la disponibilidad de software para análisis bibliométrico, y humanas, como la disponibilidad de personal capacitado para realizar los estudios. Así mismo, la apropiación y masificación del uso de estas metodologías al interior de la institución requieren de la existencia de estudios de vigilancia científica que sirvan como ejemplo o como línea base para la realización de futuros estudios aplicados en diferentes áreas de conocimiento. Sin embargo, en la actualidad no se tienen estudios realizados en la institución que tengan las características que se requieren dentro del proyecto de construcción del Plan Prospectivo para el desarrollo de Ciencia, Tecnología e Innovación. Por tal razón, se determinó por parte de los investigadores de este proyecto, la necesidad de realizar un conjunto de estudios de la investigación en un área de interés del grupo TecnIO y del grupo DYGOTOS, como es el área de la calidad del software. [5]

En este orden de ideas, el problema abordado por este proyecto es la inexistencia de estudios de tendencias y perfiles de investigación, que puedan ser utilizados como ejemplo o línea base para la realización de otros estudios de tendencias que se deben realizar en el marco del proyecto de investigación con el que se está construyendo el Plan Prospectivo para el desarrollo de Ciencia, Tecnología e Innovación de la Unimagdalena.

Teniendo en cuenta la problemática anterior, en este proyecto se desarrolló un perfil de la investigación científica en el área de conocimiento de la ingeniería del Software denominada "Calidad del Software". Esta pasantía de investigación permitió identificar las tendencias en la producción científica, los temas investigados, los actores clave (individuos, organizaciones y países) y las redes de colaboración. El cual servirá como ejemplo para la realización de perfiles de investigación en otras áreas de conocimiento de interés para los investigadores y grupos de la Universidad del Magdalena.

A continuación, se presenta, en la siguiente sección, un conjunto de teorías y definiciones utilizadas para llevar a cabo la investigación, seguido se encuentra la metodología con la que se desarrolló el proceso investigativo, luego se expone la sección de resultados, la cual se encuentra constituida por aquellos análisis de primer y segundo orden realizados durante el transcurso de la investigación, seguido de una sección de discusión donde se analizan los resultados más relevantes. Por último, se encuentra la sección de conclusiones, la cual hace referencia a los conocimientos y experiencias que se obtuvieron durante el desarrollo de la pasantía de investigación.

4. Marco teórico

En esta sección se presentan los conceptos principales que sustentan la investigación realizada.

Dado que la mira central de esta investigación estará puesta en el área del conocimiento de la ingeniería del software denominada “Cálida del software” es prescindible el abordar o plantear algunos ejes conceptuales necesarios para el entendimiento del desarrollo de este proyecto.

4.1 Ingeniería del software

El software en la actualidad es parte esencial en nuestras vidas dado que afecta a gran parte de las de las actividades que realizamos, ha invadido el comercio, la cultura y la educación. Además, que manifiesta ciertos retos a la hora de su elaboración tales como: Rendimiento, seguridad, adaptabilidad, concurrencia, disponibilidad y alta calidad entre otros más. Por tales razones es necesario el aplicar una disciplina de técnicas a los procedimientos y al desarrollo al producir un software. Es decir, aplicar ingeniería al software. Este tipo de ingeniería la define [6] como el establecimiento y uso de principios fundamentales de la ingeniería con objeto de desarrollar en forma económica software que sea confiable y que trabaje con eficiencia en máquinas reales. Entenderemos el concepto de ingería del software, del mismo modo en que el autor la ha definido.

Por su parte [7] la define como disciplina que de la ingeniería que comprende todos los aspectos de la producción de software desde las etapas iniciales de la especificación del sistema, hasta el mantenimiento de éste después de que se utiliza.

Según [5] la ingeniería del software se encuentra organizada por 15 disciplinas del conocimiento las cuales se presentan en Tabla 1.

Tabla 1 Áreas de conocimiento (SWEBOK)

Áreas de conocimiento
Requerimientos del software
Diseño del software
Construcción del software
Pruebas del software
Mantenimiento del software
Gestión de la configuración del software
Gestión en la ingeniería del software
Procesos en la ingeniera del software
Métodos y modelos de la ingeniería del software
Calidad del software
Prácticas profesionales en la ingeniería del software
Economía de la ingeniería del software

Fundamentos de computación
Fundamentos matemáticos
Fundamentos de ingeniería

Fuente: elaboración propia en base al contenido del libro SWEBOK.

4.2 Calidad del software

Al concebir la importancia del software en la vida humana, podemos inferir que el aseguramiento de la calidad del producto software es imprescindible, dado que la calidad del software es un problema frecuente en la vida diaria. Se lo tropieza de una forma u otra. Cuando en un cajero se pide una cantidad de dinero y esté entrega otra cantidad. ¿No estamos ante un problema de calidad del software?, esto nos hace entender que obtener software de alta calidad es una meta importante, según lo definido por [8] la calidad del software es la concordancia con los requisitos funcionales y de rendimiento explícitamente establecidos con los estándares de desarrollo explícitamente documentados y con las características implícitas que se espera de todo software desarrollado profesionalmente.

Puesto que la calidad de software es un aspecto ubicuo en la ingeniería del software entonces es necesario el uso de ciertos procesos para alcanzarla. Según [5] la calidad del software abarca dos subáreas fundamentales las cuales son:

- **Fundamentos de Calidad de software:** La cual comprende las definiciones básicas o los conceptos relacionados con la calidad del software.
- **Procesos de Gestión de Calidad de software:** Se refiere a las numerosas actividades aplicadas a todas las perspectivas de procesos de software, productos, y recursos para así dirigir las a satisfacer al cliente y las exigencias del personal implicado.

4.3 Bibliometría

La bibliometría es una rama de la cienciometría que estudia la producción científica aplicando métodos matemáticos y estadísticos, con el objetivo de medir el crecimiento de la ciencia por medio de la producción científicas y así poder establecer el desarrollo de las disciplinas científicas.

La bibliometría realiza los estudios de consumo de información basándose en los documentos que utilizan los científicos, es decir la producción de otros científicos [9].

5. Metodología

Para el desarrollo de este proyecto se empleó una metodología propuesta por [10]. Esta metodología se compone de nueve pasos agrupados en tres fases: inteligencia, análisis y diseño, y selección; tal como se indica en la Anexo 1. Permitiendo la identificación de



tendencias en la producción científica, los temas más investigados, los actores clave (individuos, organizaciones y países) y las redes de colaboración en el área de interés de la ingeniería del software denominada Calidad del software.

5.1 Inteligencia

En esta fase se seleccionó el enfoque de la Calidad del Software, como tema para analizar. Como frontera temporal se tomó el periodo comprendido entre los años 2011 y 2016, y el ámbito internacional se consideró como localización de estudio. Se estableció como fuente de información las bases de datos de publicaciones científicas indexadas en “SCOPUS” [3] e “ISI Web of knowledge” [4].

Seguidamente, se hizo una extracción de términos claves pertinentes al tema, con el objetivo de realizar la construcción de una ecuación de búsqueda inicial, para esto se tomó como fuente de información la *Guide to the Software Engineering* (SWEBOK) [5]. Posteriormente, se formuló la ecuación con 12 términos claves, operadores lógicos y de búsqueda. A partir de los resultados más relevantes de la primera iteración se logró introducir tres nuevos términos claves asociados al tema central dando como resultado una primera ecuación con quince términos (ver Tabla 2).

Tabla 2 Primera ecuación de búsqueda (Elaboración propia)

(TITLE-ABS(software W/0 quality) OR TITLE-ABS-KEY ((quality W/0 (validation OR fundamentals OR improvement OR assurance OR costs OR measurement OR necessity OR reviews OR models OR {Management Processes} OR characteristics OR products OR safety OR requirements OR process) AND software))) AND PUBYEAR > 2010 AND PUBYEAR < 2017

Fuente: elaboración propia.

Con base en la revisión de los resultados obtenidos en cada iteración, se agregaban o extraían términos claves con el objetivo de realizar un proceso de calibración de la ecuación. Este proceso consistió en 14 iteraciones en donde se introdujeron los nuevos términos adquiridos y se descartaron aquellos que no aportaban resultados pertinentes al objetivo de la investigación., esto con el fin de verificar que los nuevos documentos arrojados eran relevantes en cuanto al tema de Calidad del Software. Al finalizar el proceso de calibración, se obtuvo una ecuación concreta, objetiva, definida, y sesgada con todas las áreas, temáticas o disciplinas académicas permitidas por la base de datos de publicaciones científicas indexadas Scopus e ISI *Web of knowledge*. En la Tabla 3.

Tabla 3 Ecuación final para ambas bases de datos (Elaboración propia)

SCOPUS	Resultados: 3,104	ISI Web of knowledge	Resultados: 636
((TITLE-ABS (software W/0 quality) OR TITLE-ABS-KEY ((quality W/0 (improvement OR assurance OR costs OR measurement OR necessity OR reviews OR models OR management OR process OR characteristics		((TS=(software NEAR/0 quality) OR TS=((quality NEAR/0 (improvement OR assurance OR costs OR measurement OR necessity OR reviews OR models OR management	

OR products OR safety OR "Defect characterization") w/0 software)) AND NOT TITLE (proceedings OR conference OR symposium OR convention)) AND PUBYEAR > 2010 AND PUBYEAR < 2017)	OR process OR characteristics OR products OR safety OR "Defect characterization") NEAR/0 software))) NOT TI= (proceedings OR conference OR symposium OR convention)) AND PY=(2011-2016)
Fecha: 29 de abril 2017 4:45 pm	Fecha: 29 de abril 2017 5:10 pm

Fuente: elaboración propia.

Posteriormente, se le aplicó a la ecuación final una prueba de distribución normal (Campana de Gauss), para corroborar la calidad de la ecuación, es decir, verificar si los resultados arrojados por la ecuación de búsqueda pertenecían al tema de análisis. Para realizar la prueba se tomó una muestra, la cual fue calculada utilizando la fórmula que se presenta en la Ecuación 1. El tamaño de la muestra fue de 341 (n=341) elementos. La selección de cada elemento de la muestra se realizó aleatoriamente con los resultados arrojados por la ecuación de búsqueda final en la base de datos Scopus.

Luego se realizó una evaluación minuciosa de cada uno de los artículos y se determinó si estaba relacionado con el área de la calidad del software. Esta evaluación se realizó revisando el título y el resumen de cada uno de los documentos de acuerdo con los conocimientos previos obtenidos durante la construcción del marco teórico, descartando así aquellos que no se consideraban pertinentes. Al terminar de examinar los artículos obtenidos, se llegó a aceptar la ecuación con un porcentaje de error del 21.021% el cual fue discutido con el tutor experto en el tema, determinando que la ecuación de búsqueda final tenía un porcentaje de error aceptable para este tipo de ejercicios.

Ecuación 1. Fórmula para el cálculo de una muestra

Prueba de distribución normal para calcular la muestra

$$n = \frac{N \cdot Z^2 \cdot p \cdot (1-p)}{(N-1) \cdot e^2 + Z^2 \cdot p \cdot (1-p)}$$

N: 3,104	→ Tamaño total de la población.
Z: 1,96	⇒ Desviación del valor medio.
E: 5%	⇒ Margen de error máximo.
P: 50%	⇒ Porción que se espera encontrar.
n	→ Tamaño de la muestra.

5.2 Análisis y diseño

En esta etapa se realizó un proceso de depuración de la ecuación, eliminando redundancias y variaciones innecesarias en los resultados finales. Esto se logró a través de la herramienta para análisis bibliométrico HAB [11]. Esta herramienta permite el procesamiento de los metadatos de cada uno de los documentos, los cuales fueron importados con el fin de normalizar la información reduciendo registros duplicados y la búsqueda de firmas de afiliaciones, revistas y autores similares. Una vez se realizó el proceso de limpieza se obtuvo como resultado 3101 documentos de Scopus y 635 documentos de ISI web. Finalmente se exportaron los metadatos en un formato CSV para su respectivo análisis en la herramienta Sci2 [12].

Una vez los metadatos fueron exportados en la herramienta, se inició un análisis de primer orden, el cual permitió realizar un análisis estadístico exploratorio de los datos. Los resultados obtenidos, fueron listas y tablas de frecuencias, variaciones de producciones científicas, temas más trabajados, el volumen de la producción científica (individual, organizacional y nacional) y el impacto científico (citaciones). Seguido de un análisis de segundo orden que permitió la construcción de matrices de concurrencia y tablas de contingencia para identificar redes de colaboración entre autores, afiliaciones y redes temáticas.

5.3 Selección

Finalmente, en la fase de selección, se realizó la representación de los diferentes análisis de primer y segundo orden, mediante el módulo de visualización de la herramienta Sci2. A continuación se muestran los resultados obtenidos (ver Tabla 4).

Tabla 4 Resumen de resultados.

Fase	Datos iniciales	Datos finales	Descripción
Inteligencia	5.403	3,104	Durante esta fase se desarrolló la primera ecuación de búsqueda con un total 5,403 resultados, luego se inició un proceso de calibración obteniendo una ecuación final con 3,104 resultados.
Análisis y diseño	3,104	3,101	Durante esta fase se depuro los resultados obtenidos con la ecuación de búsqueda empleando las herramientas HAB [11] y Sci2 [12].

Fuente: elaboración propia.

6. Resultados

En esta sección se van a presentar los resultados obtenidos en la investigación la cual abarca el periodo comprendido entre los años 2011-2016.

6.1 Análisis de primer orden

En este apartado se realiza una representación descriptiva y gráfica de los análisis acordados para esta sección, de dichos análisis se tiene: el número de documentos, su distinción por tipo de documentos en el tiempo, la producción del top 10 por países en el tiempo, producción del top 10 de afiliaciones y autores en el tiempo, y los documentos más citados entre otros. Tomando como referencia los resultados arrojados por las bases de datos Scopus e ISI web, los datos necesarios para exponer los análisis anteriormente mencionados.

6.1.1 Documentos por año.

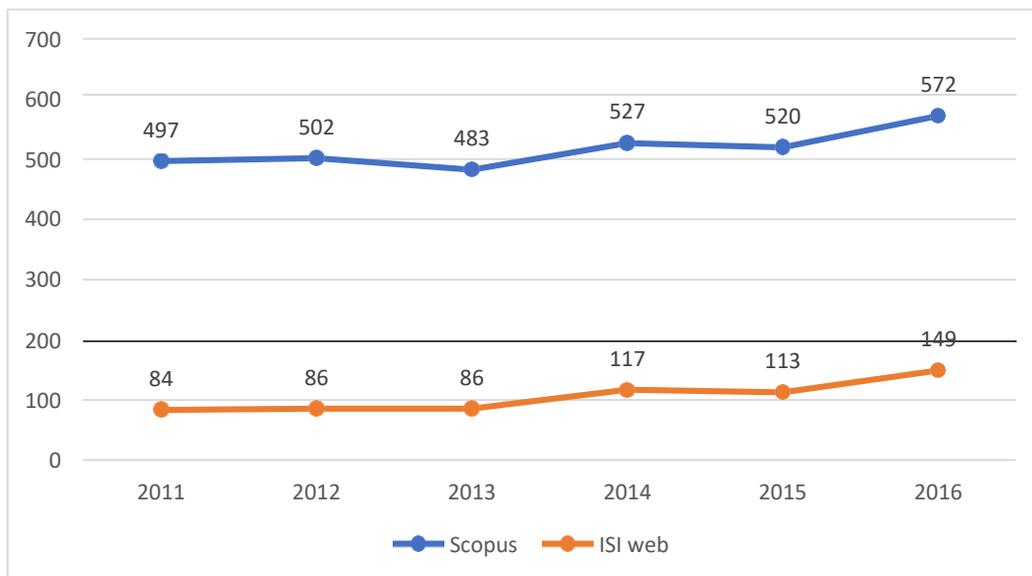
En este análisis se observa la cantidad de documentos analizados en función del tiempo (ver Tabla 5), dichos documentos fueron extraídos de las bases de datos indexadas Scopus e ISI Web. De acuerdo al comportamiento que se aprecia en la Ilustración 1, se puede inferir que para la base de datos ISI Web la cantidad de documentos se mantiene una tendencia incremental a medida que transcurre los años, en cambio para Scopus es variante. También, se puede apreciar que para Scopus el año donde menos se identificaron artículos publicados fue en el 2013, y para ISI Web fue el 2011, en contra parte el año 2016 fue donde se identificaron más presencia de artículos publicados para ambas bases de datos.

Tabla 5 Número de documentos por año (Scopus)

Años	Scopus	ISI web
2011	497	84
2012	502	86
2013	483	86
2014	527	117
2015	520	113
2016	572	149
Total	3101	635

Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

Ilustración 1 Cantidades de documentos por año



Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

6.1.2 Producción de documentos por tipo

En este análisis se determinó la producción de los documentos seleccionados para el estudio, considerando su tipo de documento por base de datos, de igual forma muestra el comportamiento de dichos tipos de documentos durante el periodo de tiempo analizado.

De los resultados obtenidos se puede inferir que para la base de datos Scopus el tipo de documento más significativo es el de artículo publicado en conferencia científica, el cual equivale al 65.4% del total de los documentos obtenidos por esta base de datos. De igual forma, el tipo de documento artículo publicado en revista indexada también es relevante dado que se identificó que el 28.7% del total de los documentos, Así mismo, se identificaron artículos en espera de ser publicados los cuales constituyen el 1% del total de documentos obtenidos por la base de datos Scopus (ver Tabla 6). En la Ilustración 2 se puede apreciar gráficamente la distribución descrita.

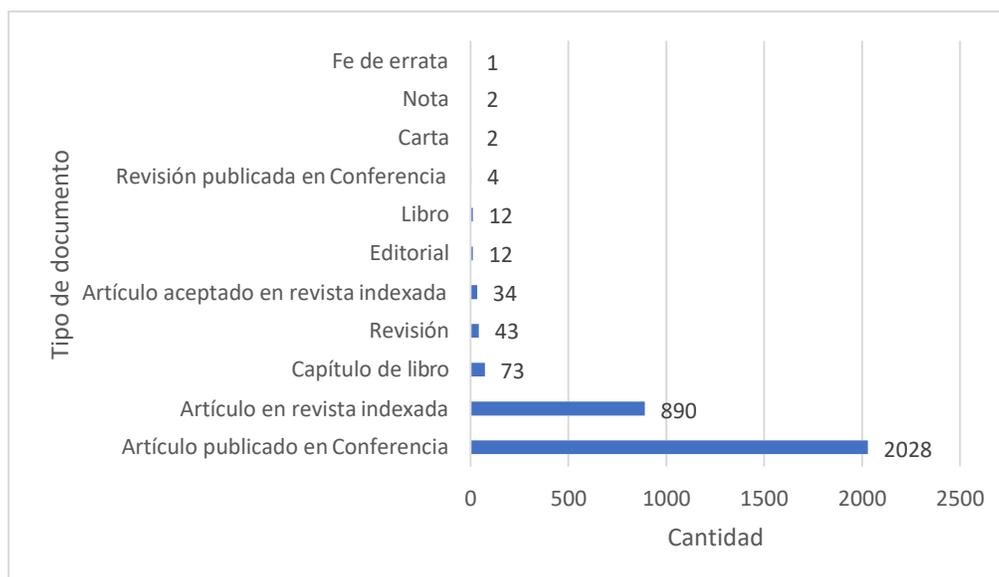
En cuanto al comportamiento al transcurrir los años se observa el tipo de documento artículo en revista indexada ha tenido un crecimiento al transcurrir los años esto se puede apreciar en la Ilustración 3.

Tabla 6 Distribución de los documentos en estudio, publicados en Scopus por tipo y por año

Tipos documentos	2011	2012	2013	2014	2015	2016	Total
Artículo publicado en Conferencia	350	357	315	325	333	348	2028
Artículo en revista indexada	127	127	136	170	153	177	890
Capítulo de libro	5	4	25	15	17	7	73
Revisión	10	10	2	3	6	12	43
Artículo en prensa	0	0	0	2	5	27	34
Editorial	1	2	1	3	4	1	12
Libro	2	1	2	5	2	0	12
Revisión publicada en Conferencia	0	0	2	2	0	0	4
Carta	1	0	0	1	0	0	2
Nota	0	1	0	1	0	0	2
Fe de errata	1	0	0	0	0	0	1
Total	497	502	483	527	520	572	3101

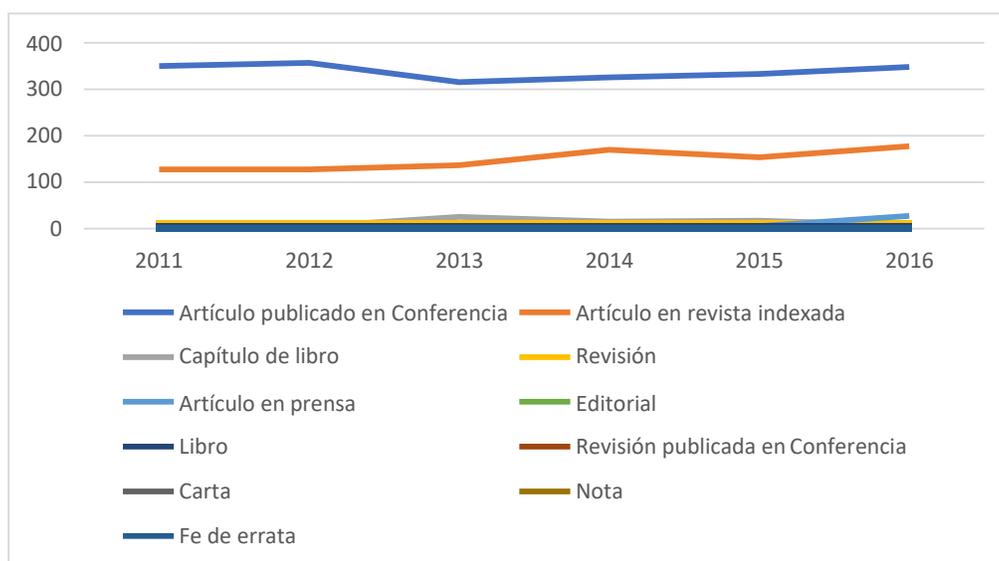
Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

Ilustración 2 Distribución de los documentos en estudio, publicados en Scopus por tipo



Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

Ilustración 3 Distribución de los documentos en estudio, publicados en Scopus por tipo y por año



Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

Por su parte, en Ilustración 4 se deduce que para la base de datos ISI Web el tipo de documento más utilizado para las investigaciones o publicaciones acerca del tema de la calidad del software es el tipo de documento artículo en revista indexada el cual abarca el 91% del total de documentos obtenidos por ésta base de datos. La afirmación presentada se apoya en los datos en evaluación que conforman la Tabla 7.

En cuanto al comportamiento al transcurrir los años se observa que para la base de datos ISI Web el tipo de documento artículo en revista indexada ha tenido un crecimiento al transcurrir los años esto se puede apreciar en la Ilustración 5.

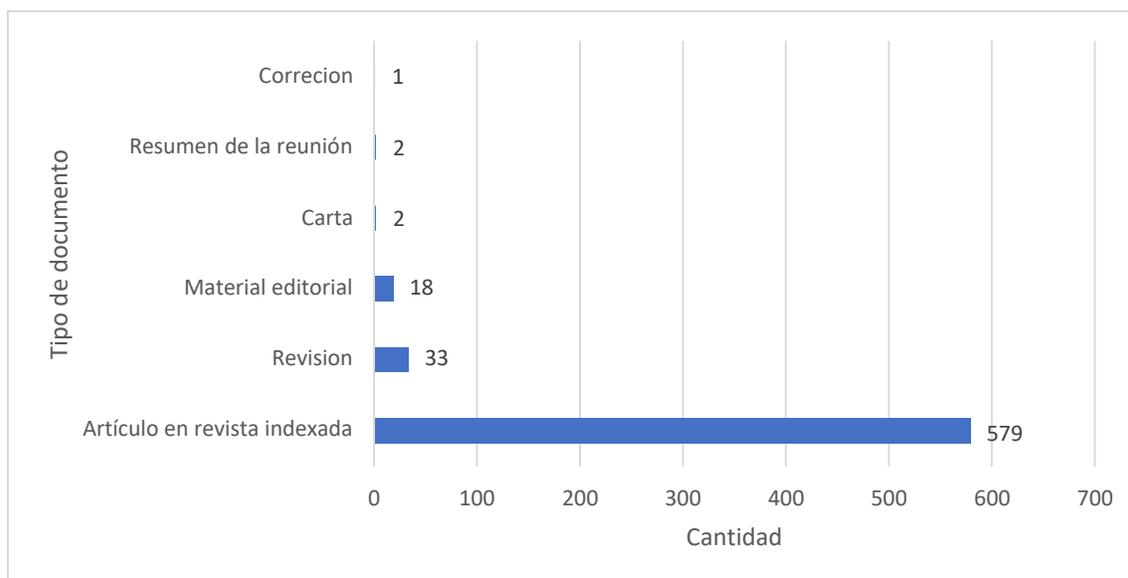
Tabla 7 Distribución de los documentos en estudio, publicados en ISI Web por tipo y por año

Tipos documentos	2011	2012	2013	2014	2015	2016	Total
Artículo en revista indexada	76	79	81	112	98	133	579
Revisión	5	5	2	4	8	9	33
Material editorial	1	2	3	0	6	6	18
Carta	1	0	0	1	0	0	2
Resumen de la reunión	0	0	0	0	1	1	2
Corrección	1	0	0	0	0	0	1
Total	84	86	86	117	113	149	635

Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

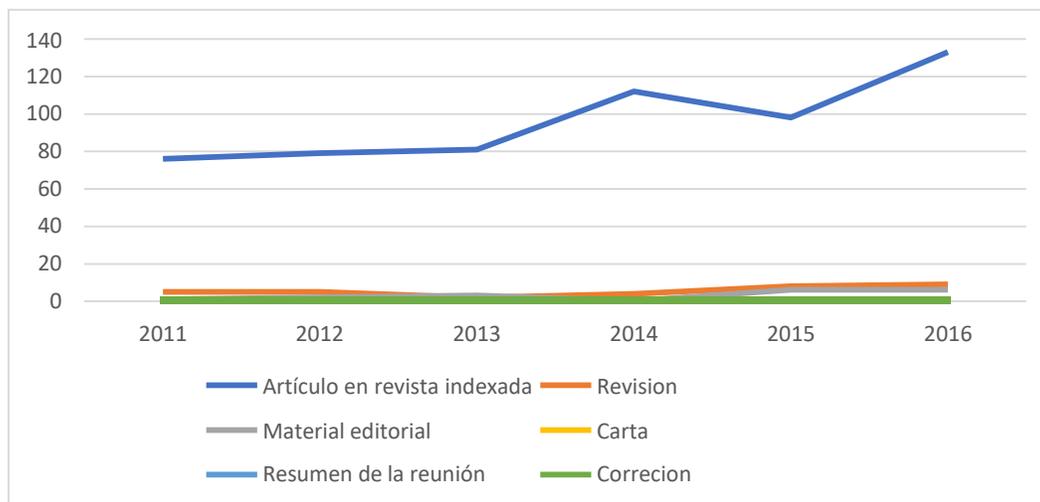


Ilustración 4 Distribución de los documentos en estudio, publicados en ISI Web por tipo



Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

Ilustración 5 Distribución de los documentos en estudio, publicados en ISI Web por tipo y por año



Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

6.1.3 Top 10 de instituciones con la mayor producción (Scopus)

En este análisis se determinó las instituciones con el mayor número de publicaciones científicas extraídas de los resultados obtenidos por la base de datos Scopus. Permitiendo, identificar tendencias de las instituciones en cuanto a la producción de documentos científicos (ver Tabla 8). Así mismo, se determinó que Florida Atlantic University produjo el 1.48% del total de documentos obtenidos, Eastern Connecticut State University el 1.1% y Queen's University con él 1% siendo estas 3 las afiliaciones con más documentos vinculados. En contra parte se identificó que las instituciones University of California y University of Szeged las cuales comparten la misma cantidad de producciones científicas quedaron él lo último del escalafón con 0.5% del total de los documentos. En la Ilustración 6 e Ilustración 7 se presentan gráficamente los resultados obtenidos en este análisis.

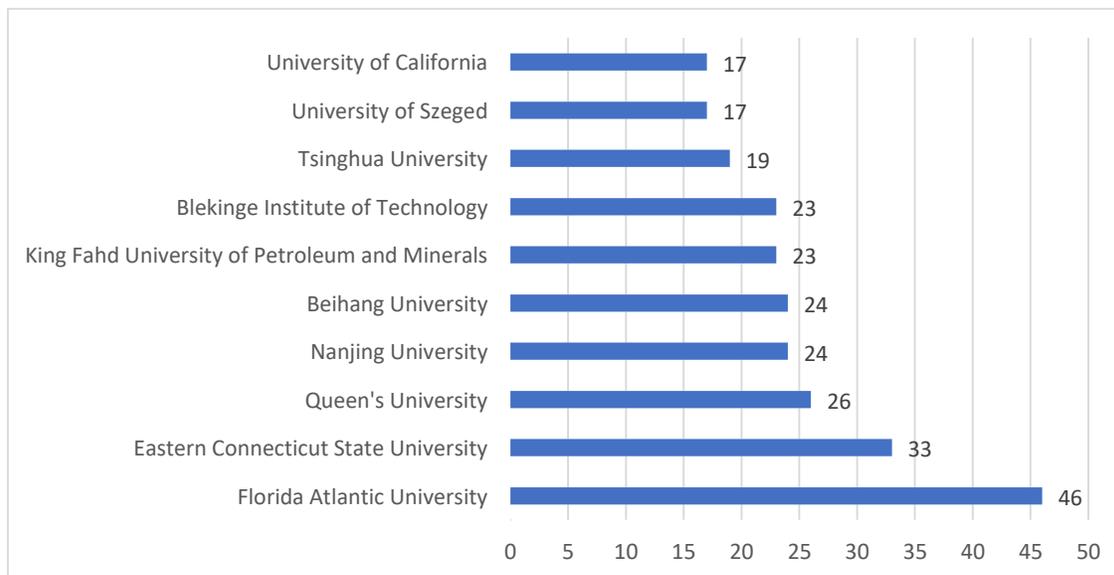
Es importante decir, que este análisis no se pudo realizar con los resultados obtenidos en la base de datos ISI Web, ya que la misma no proporciona datos de las instituciones al momento de realizar la exportación de datos.

Tabla 8 Top 10 de instituciones con la mayor producción de los documentos en estudio, publicados en Scopus, por año

Afiliación	2011	2012	2013	2014	2015	2016	Total
Florida Atlantic University	9	8	8	10	9	2	46
Eastern Connecticut State University	7	7	4	7	7	1	33
Queen's University	5	4	3	3	7	4	26
Nanjing University	5	3	2	3	5	6	24
Beihang University	2	5	2	4	8	3	24
King Fahd University of Petroleum and Minerals	7	2	8	2	2	2	23
Blekinge Institute of Technology	5	3	2	5	1	7	23
Tsinghua University	2	4	8	4	1	0	19
University of Szeged	2	5	1	4	5	0	17
University of California	7	2	2	5	1	0	17
Total	51	43	40	47	46	25	252

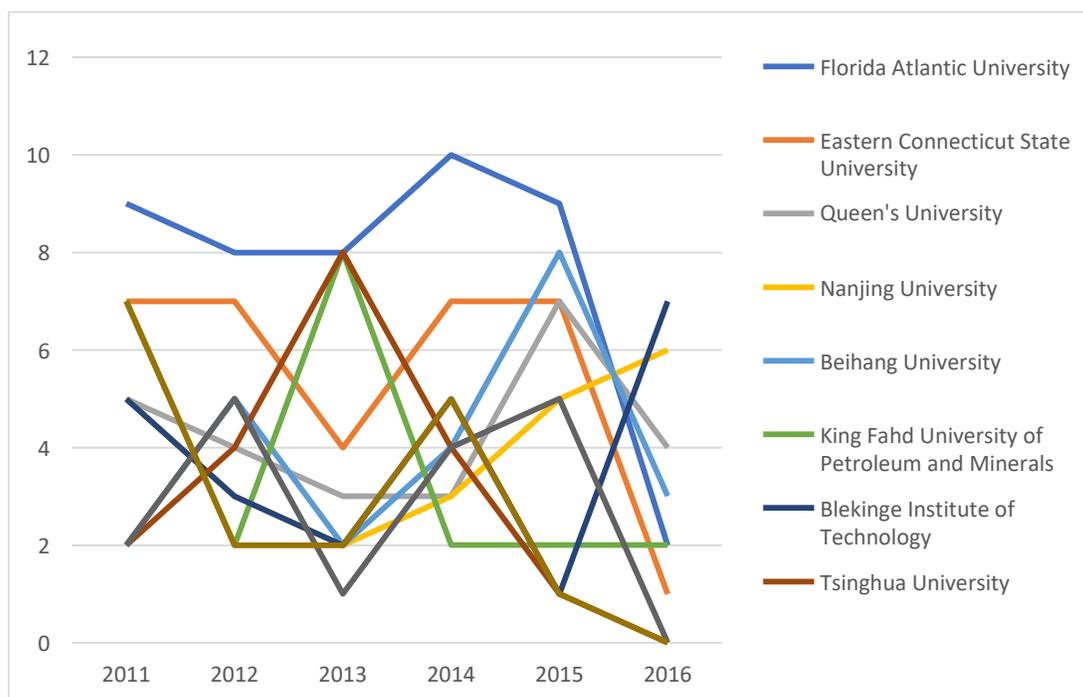
Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

Ilustración 6 Top 10: Top 10 de instituciones con la mayor producción de los documentos en estudio, publicados e Scopus, por año



Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

Ilustración 7 Distribución de los documentos publicados por el Top 10 de instituciones con la mayor producción de documentos en estudio, publicados en Scopus, por año



Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

6.1.4 Top 10 de autores con mayor producción de documentos seleccionados para el análisis

En este análisis se observa el top 10 de los autores con más publicaciones científicas durante el periodo de análisis. Se encontró que para los documentos seleccionados en la base de datos Scopus e ISI Web, Khoshgoftaar T.M es el autor con mayor producción científica el cual comprende un 1.48% de un total de 3101 documentos en Scopus (ver Tabla 9) y un 1.9% de un total de 635 documentos en ISI Web (ver Tabla 10).

Así mismo, se demuestra que, el autor Hassan AE comparte el podio con las mismas producciones científicas en la base de datos ISI Web pero su resultado es muy diferente en la base de datos Scopus dando ver que se encuentra por debajo del 1% en cuanto a producciones científicas. Esta diferencia se puede apreciar gráficamente en la Ilustración 8 para Scopus e Ilustración 9 para ISI Web.

Por otra parte, el autor Gao K es muy relevante para Scopus comprendiendo el 1% de un total de 3101 documentos, el resto de autores para esta misma base de datos se mantuvieron debajo de este porcentaje. Además, se observa que la producción científica de los autores en la base de datos Scopus se vio afectada en los años 2013 y 2016 la cual tuvo una tendencia a decrementar (ver Ilustración 10), caso contrario a la base de datos ISI Web donde el 2016 fue uno de los más relevante y su tendencia en la línea del tiempo fue incremental (ver Ilustración 11).

Tabla 9 Top 10 de autores con mayor producción de documentos seleccionados publicados en Scopus, por año

Producción de documentos Autor	2011	2012	2013	2014	2015	2016	Total
Khoshgoftaar T.M	9	9	8	10	9	1	46
Gao K	7	7	4	7	7	1	33
Napolitano A	2	5	4	4	8	1	24
Wang H	6	3	5	3	2	2	21
Hassan A.E	2	2	3	3	5	6	21
Zhang Y	4	5	2	5	2	0	18
Piattini M	3	2	2	4	5	1	17
Liu S	2	3	2	3	2	5	17
Zhang H	1	4	4	5	1	1	16
Malhotra R	0	1	2	5	3	4	15
Total	36	41	36	49	44	22	228

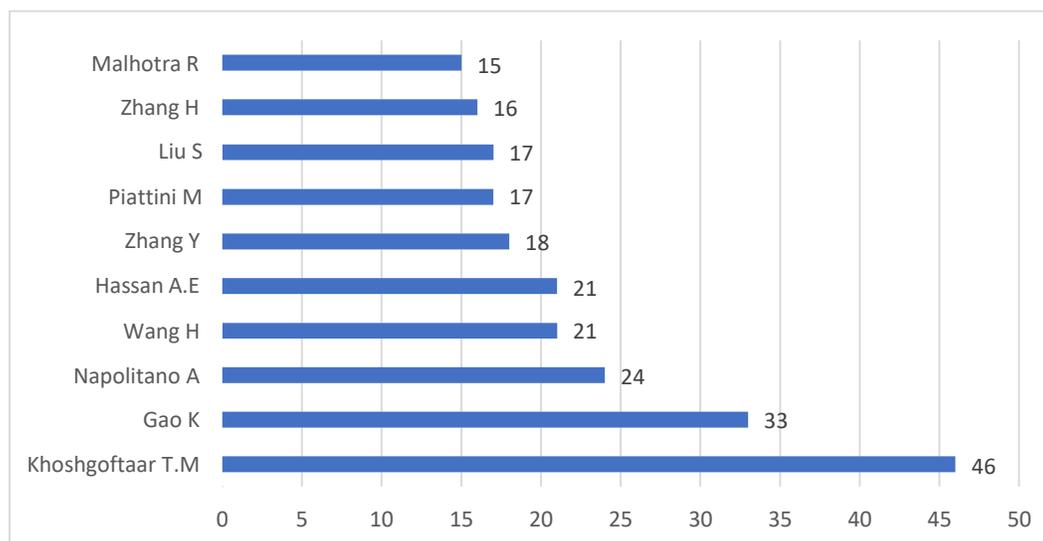
Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

Tabla 10 Top 10 de autores con mayor producción de documentos seleccionados publicados en ISI Web, por año

Producción de documentos de Autor	2011	2012	2013	2014	2015	2016	Total
Khoshgoftaar, TM	3	2	0	3	4	0	12
Hassan, AE	1	1	2	1	4	3	12
Lee, S	2	1	0	3	2	2	10
Gao, KH	2	2	0	2	2	0	8
Adams, B	1	1	1	0	3	2	8
Zhang, L	1	1	0	2	0	3	7
Misra, S	2	1	0	1	0	3	7
Liu, H	1	0	2	1	0	3	7
Kessentini, M	0	0	1	3	2	1	7
Alshayeb, M	2	0	2	1	2	0	7
Total	15	9	8	17	19	17	85

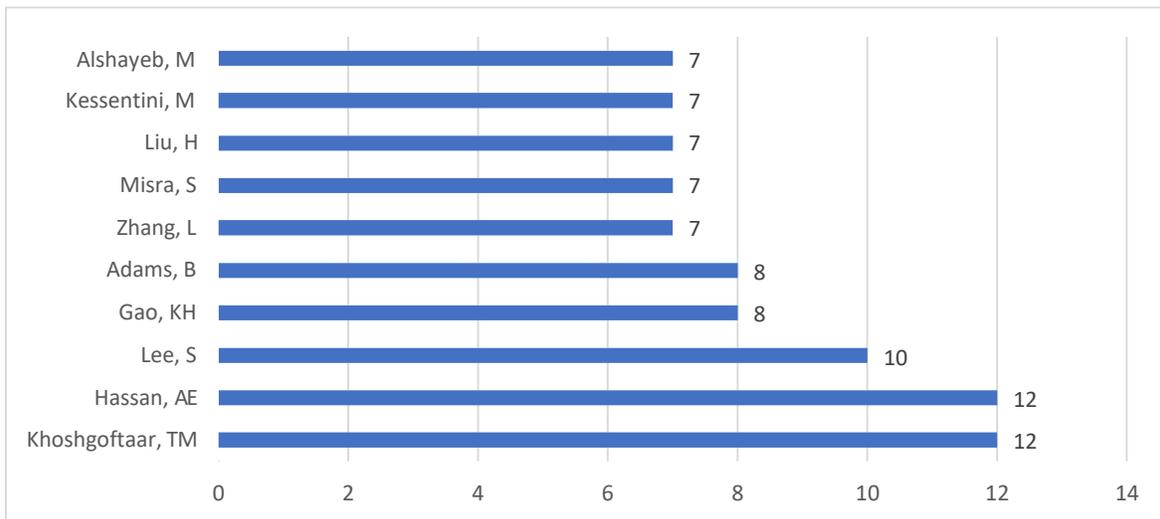
Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

Ilustración 8 Distribución del Top 10 de autores con mayor producción de documentos seleccionados publicados en Scopus



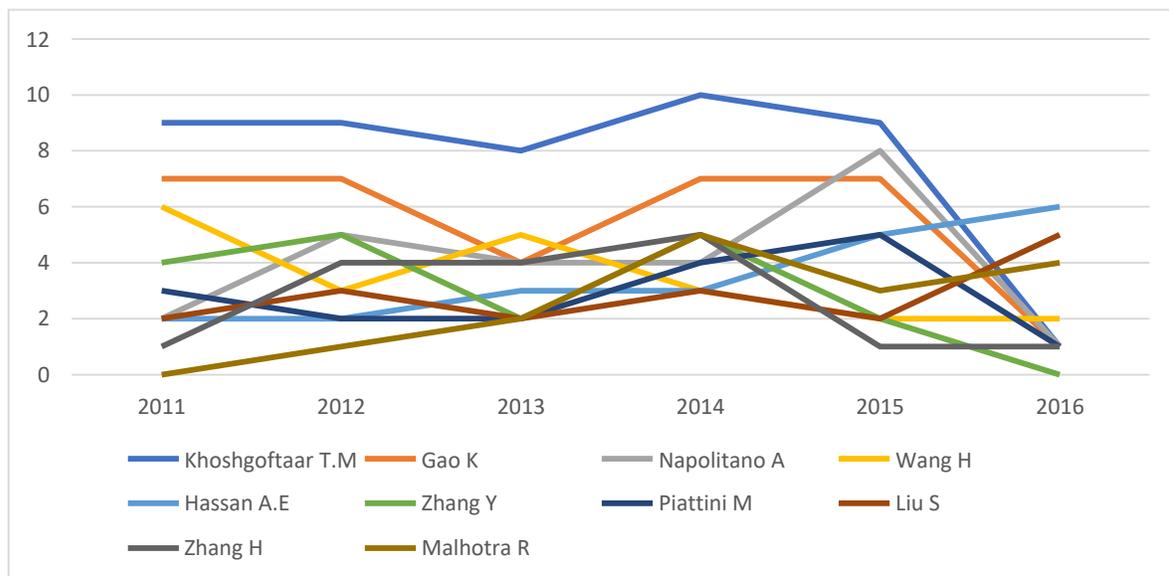
Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda

Ilustración 9 Distribución del Top 10 de autores con mayor producción de documentos seleccionados publicados en ISI Web



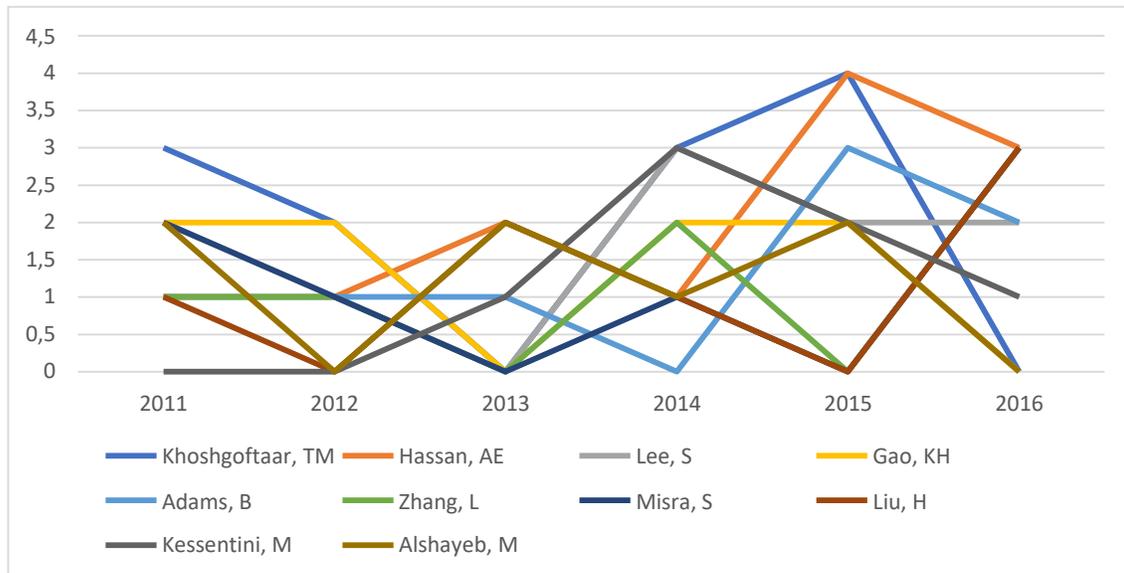
Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda

Ilustración 10 Distribución del Top 10 de autores con mayor producción de documentos seleccionados publicados en Scopus, por año



Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda

Ilustración 11 Distribución del Top 10 de autores con mayor producción de documentos seleccionados publicados en ISI Web, por año



Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda

6.1.5 Top 10 de documentos por países(Scopus)

En este análisis se puede observar la producción de documentos científicos agrupados según los países de producción, para los resultados extraídos de la base de datos Scopus, se evidencia que los países con mayor número de publicaciones científicas son India con 11.8%, seguido de Estados Unidos con un 11% y China con 10.9% de un total de 3101 documentos extraídos, mientras que la producción de otros países se encuentra por debajo del 5%, esto se puede apreciar gráficamente en la Ilustración 12. Por otra parte, en la Tabla 11 se demuestra que China a través de los años de análisis decayó en la producción de documentos en los años 2015 y 2016, a diferencia de la India y Estados Unidos quienes se mantuvieron estables (ver Ilustración 13).

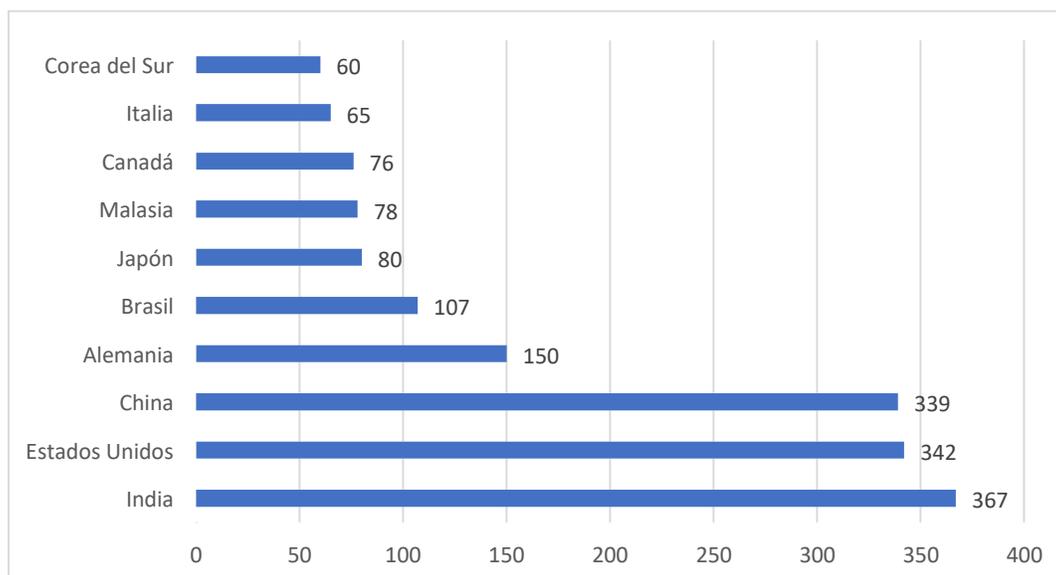
Es importante decir, que este análisis no se pudo realizar con los resultados obtenidos en la base de datos ISI Web, ya que la misma no proporciona datos de países al momento de realizar la exportación de datos.

Tabla 11 Top 10 de los países con mayor producción de documentos seleccionados publicados en Scopus, por año

Países	2011	2012	2013	2014	2015	2016	Total
India	46	40	45	61	93	82	367
Estados Unidos	65	60	58	52	54	53	342
China	61	78	56	79	26	39	339
Alemania	34	31	29	20	15	21	150
Brasil	15	10	17	14	21	30	107
Japón	11	19	17	8	12	13	80
Malasia	17	9	13	15	12	12	78
Canadá	8	16	16	6	16	14	76
Italia	5	11	13	13	10	13	65
Corea del Sur	3	13	2	15	13	14	60
Total	265	287	266	283	272	291	1664

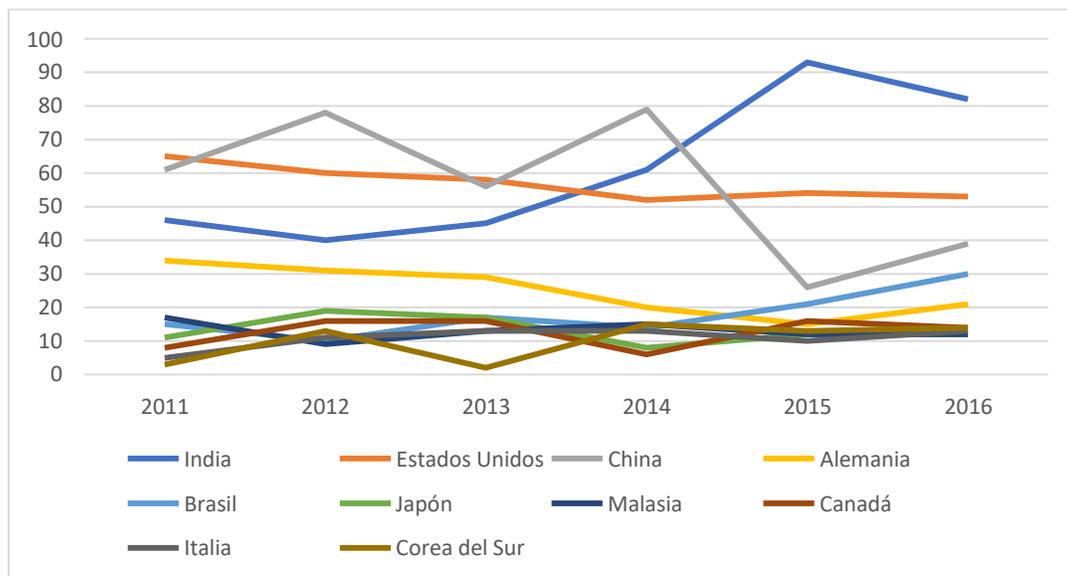
Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

Ilustración 12 Distribución del Top 10 de los países con mayor producción de documentos seleccionados publicados en Scopus



Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

Ilustración 13 Top 10 de los países con mayor producción de documentos seleccionados publicados en Scopus, por año



Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

6.1.6 Top 10 de los documentos más citados (Scopus)

En este análisis se presentan los documentos extraídos de la base de datos Scopus con mayor número de citas, debido a la gran cantidad de documentos resultantes, se muestra un top 10 de los documentos con mayor impacto científico. Como se puede observar en la Tabla 12, en donde se presentan los documentos más importantes en marco de la calidad el software para la realización de proyectos de investigación, se demuestra que el documento *Carrots and rainbows: Motivation and social practice in open source* es el más citado e importante de la investigación por la tendencia a ser utilizado como referente para realizar proyectos o documentos de investigación en el área de la calidad del software. En la Ilustración 14 se presenta gráficamente los resultados obtenidos en este análisis.

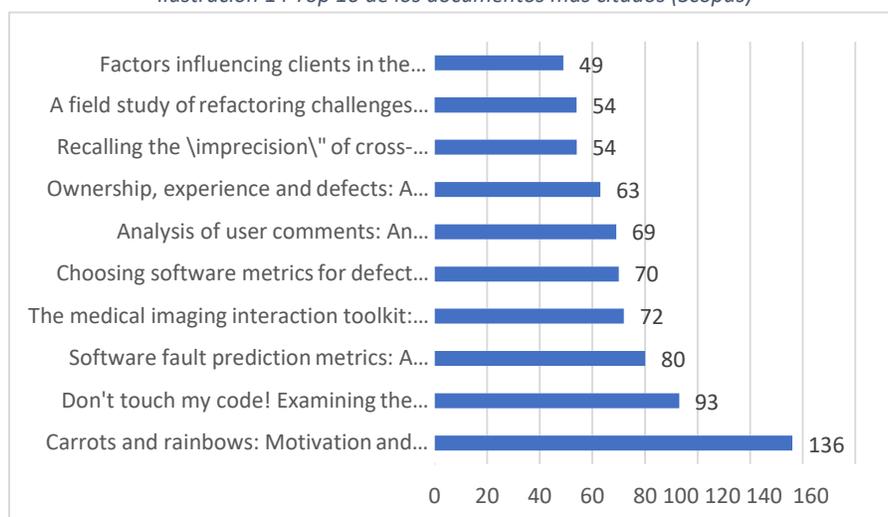
Es importante decir, que este análisis no se pudo realizar con los resultados obtenidos en la base de datos ISI Web, ya que la misma no proporciona datos de citas.

Tabla 12 Top 10 de los documentos más citados (Scopus)

Documento	Citaciones
<i>Carrots and rainbows: Motivation and social practice in open source software development [13]</i>	136
<i>Don't touch my code! Examining the effects of ownership on software quality [14]</i>	93
<i>Software fault prediction metrics: A systematic literature review [15]</i>	80
<i>The medical imaging interaction toolkit: Challenges and advances: 10 years of open-source development [16]</i>	72
<i>Choosing software metrics for defect prediction: An investigation on feature selection techniques [17]</i>	70
<i>Analysis of user comments: An approach for software requirements evolution [18]</i>	69
<i>Ownership, experience and defects: A fine-grained study of authorship [19]</i>	63
<i>Recalling the imprecision "of cross-project defect prediction" [20]</i>	54
<i>A field study of refactoring challenges and benefits [21]</i>	54
<i>Factors influencing clients in the selection of offshore software outsourcing vendors: An exploratory study using a systematic literature review [22]</i>	49

Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

Ilustración 14 Top 10 de los documentos más citados (Scopus)



Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

6.1.7 Top 10 de las palabras claves más utilizadas(Scopus)

En este análisis se observa las palabras claves más utilizadas en los documentos extraídos de la base de datos Scopus. En la Ilustración 15 se evidencia que los términos claves más influyentes o significativos en marco de la calidad del software son: *Software quality*, *Software engineering*, *software metrics*.

Por otra parte, se demuestra en la Ilustración 16 que el termino *Software quality* presentó una decadencia en su uso en el año 2013. De igual forma se aprecia que en los años 2015 y 2016 ha habido un aumento en el uso de los términos claves que comprenden el top. En la Tabla 13 se aprecia el listado de palabras clave más utilizadas en los documentos extraídos de la base de datos Scopus en marco al área del conocimiento calidad del software.

Este análisis no se pudo realizar con los resultados obtenidos en la base de datos ISI Web, ya que la misma no proporciona estos datos al momento de realizar la exportación.

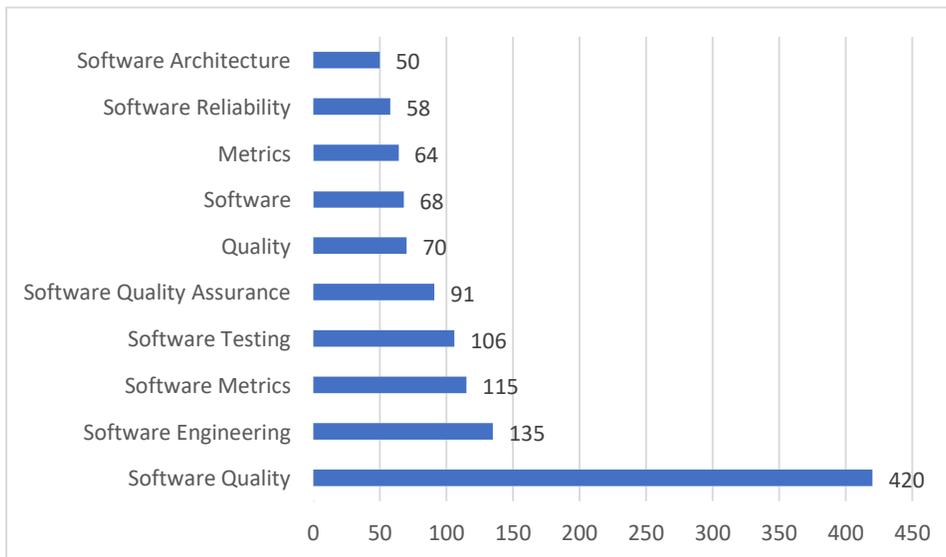
Tabla 13 Top 10 de las palabras claves más utilizadas (Scopus)

Palabras claves	2011	2012	2013	2014	2015	2016	Total
Software Quality	65	76	52	69	74	84	420
Software Engineering	17	16	26	17	28	31	135
software Metrics	20	22	18	12	24	19	115
Software Testing	14	14	13	22	17	26	106
Software Quality Assurance	11	14	12	14	14	26	91
Quality	12	16	3	10	12	17	70
Software	9	7	7	11	19	15	68
Metrics	5	11	8	9	18	13	64
Software Reliability	5	7	12	11	8	15	58
Software Architecture	10	6	5	10	12	7	50
Total	168	189	156	185	226	253	1177

Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

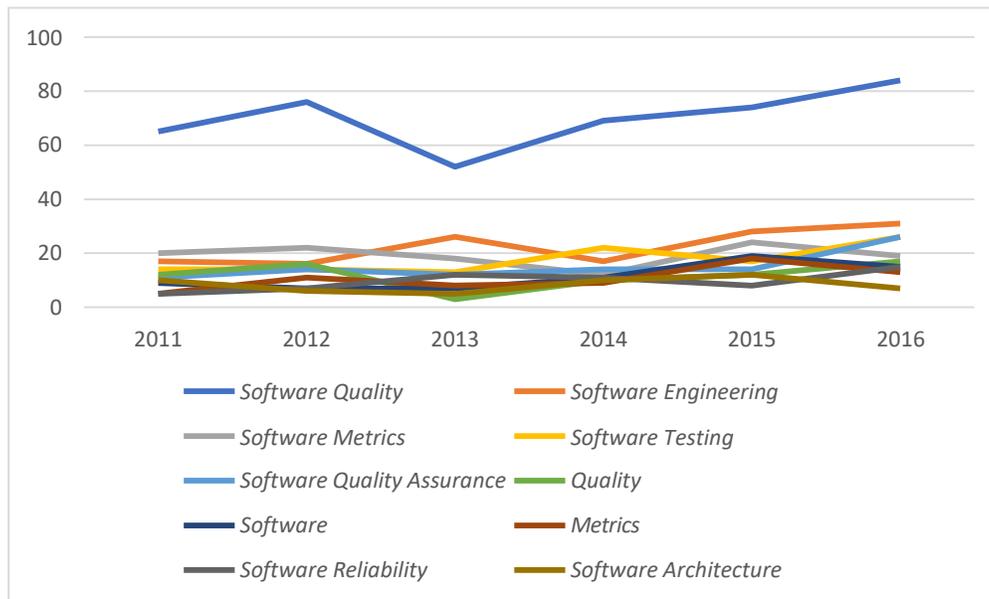


Ilustración 15 Top 10 de las palabras claves más utilizadas (Scopus)



Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

Ilustración 16 Top 10 de las palabras claves más utilizadas por año (Scopus)



Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

6.1.8 Top 10 de las palabras más utilizadas de la bolsa de palabras

En este análisis se puede apreciar el top 10 de las palabras más recurrentes extraídas de la bolsa de palabras, podemos observar que los términos más significativos o utilizados en Scopus y en ISI Web son: *software quality* y *software*, esto puede verse reflejado gráficamente en la Ilustración 17 e Ilustración 18. Así mismo se evidencia que

ambas bases de datos comparten cierta similitud en los términos que comprenden el top de sus respectivos análisis (ver Tabla 14 para Scopus y Tabla 15 para ISI Web).

Por otra parte, en la Ilustración 19 Scopus e Ilustración 20 ISI Web es donde se observa el comportamiento de estos términos con respecto al tiempo, se puede apreciar que: la usabilidad del término *Software Quality* fue incremental en los 3 últimos años para ambas bases de datos. Cosa que no paso con el término “*software*” su tendencia fue incremental para la base de datos ISI Web, pero para Scopus fue decremental.

Tabla 14 Top 10 de las palabras más utilizadas de la bolsa de palabras (Scopus)

Bolsa palabras	2011	2012	2013	2014	2015	2016	Total
Software Quality	251	243	209	233	242	257	1435
Software	135	101	120	130	168	154	808
Software Development	76	70	64	87	77	63	437
Software Engineering	56	29	37	49	56	55	282
Software Systems	49	37	30	31	39	31	217
Software Quality Assurance	23	28	30	21	30	31	163
Software Product	24	21	30	25	29	32	161
Source Code	27	18	26	26	24	37	158
Software Metrics	22	31	18	20	29	30	150
Quality Software	29	20	16	28	32	22	147
Total	692	598	580	650	726	712	3958

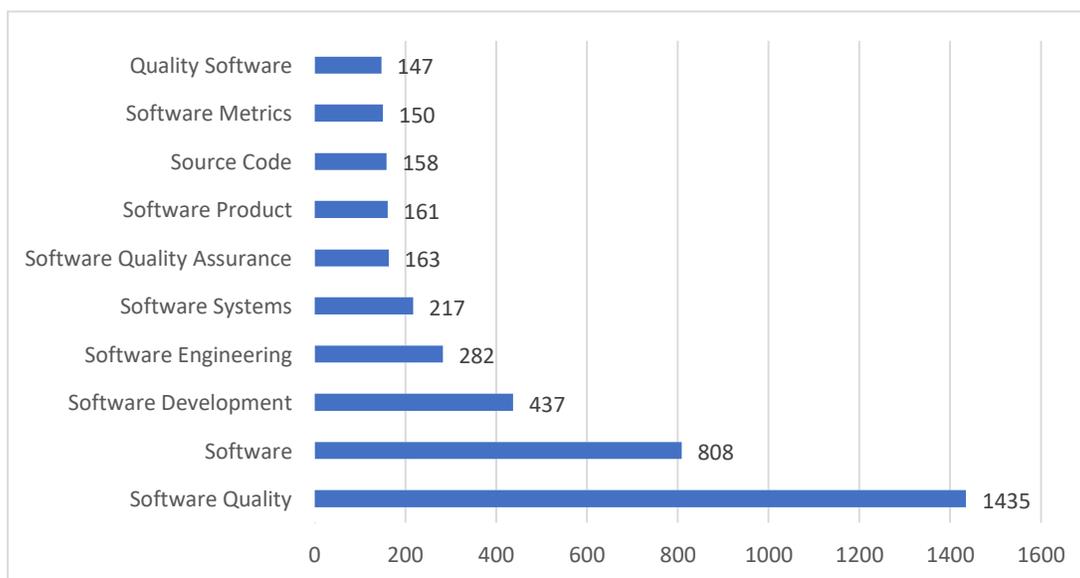
Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

Tabla 15 Top 10 de las palabras más utilizadas de la bolsa de palabras (ISI Web)

Bolsa palabras	2011	2012	2013	2014	2015	2016	Total
Software Quality	31	31	33	38	42	52	227
Software	18	18	19	24	23	47	149
Software Development	15	18	8	20	22	17	100
Software Engineering	10	4	9	11	6	8	48
Software Systems	9	6	6	7	8	10	46
Source Code	1	1	13	6	4	8	33
Software Metrics	7	5	4	3	9	4	32
Test Cases	7	10	2	3	4	5	31
Quality	4	3	4	8	4	5	28
Software Projects	5	7	0	2	7	6	27
Total	107	103	98	122	129	162	721

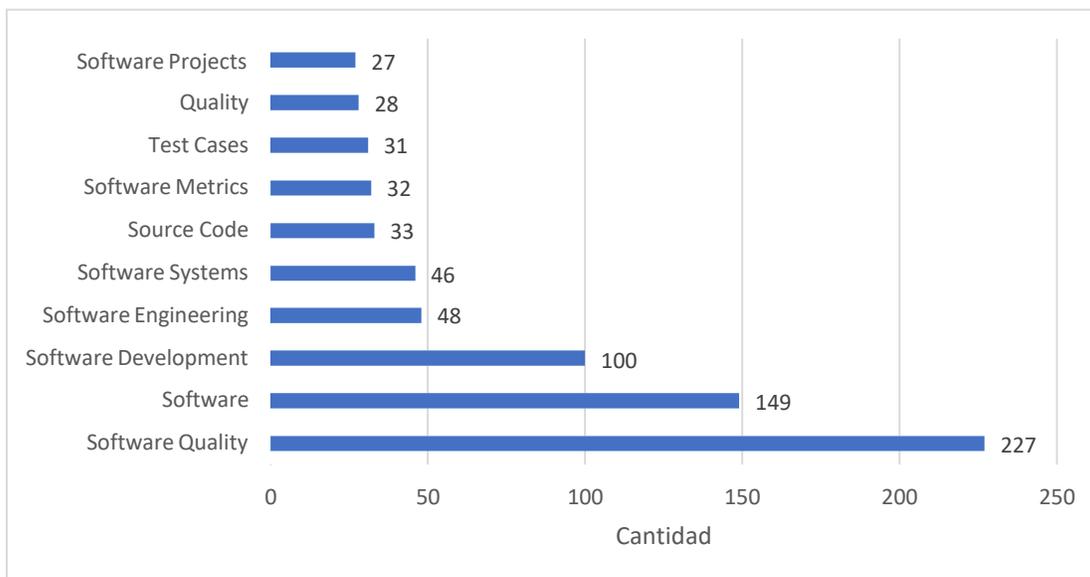
Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

Ilustración 17 Top 10 de las palabras más utilizadas de la bolsa de palabras (Scopus)



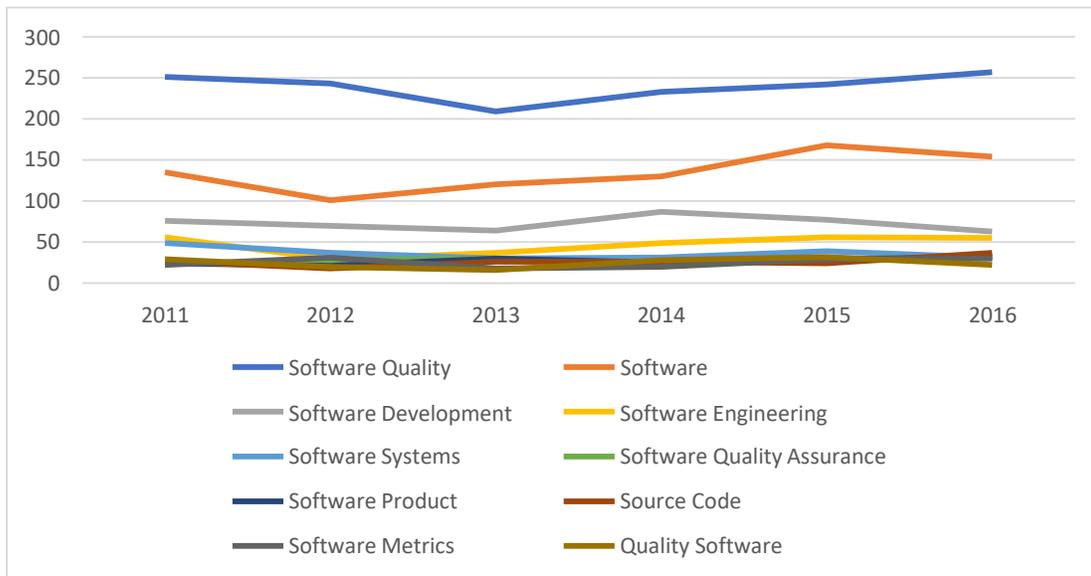
Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

Ilustración 18 Top 10 de las palabras más utilizadas de la bolsa de palabras (ISI Web)



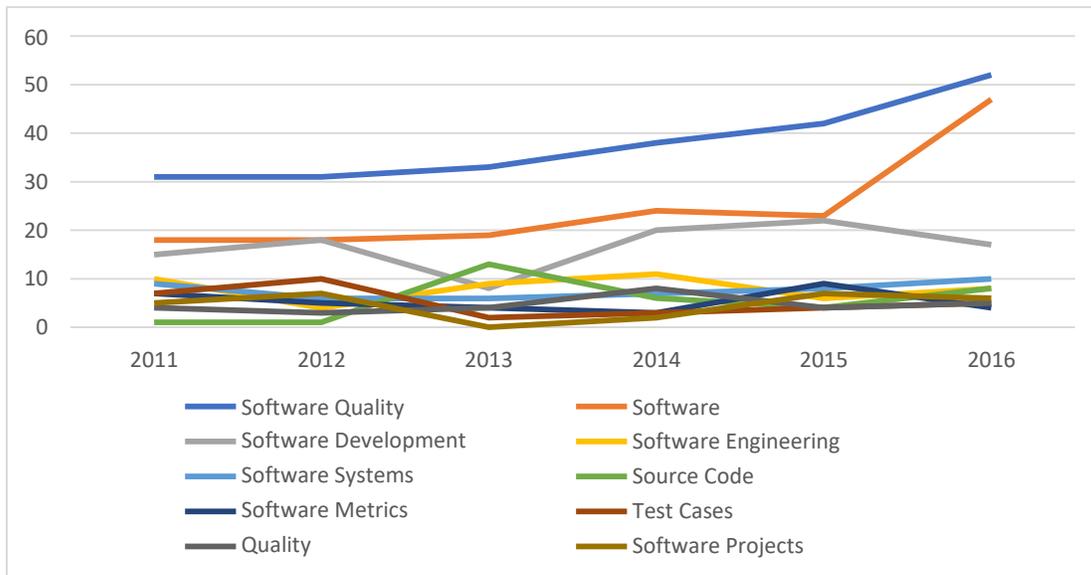
Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

Ilustración 19 Top 10 de las palabras más utilizadas de la bolsa de palabras en los años (Scopus)



Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

Ilustración 20 Top 10 de las palabras más utilizadas de la bolsa de palabras en los años (ISI Web)



Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

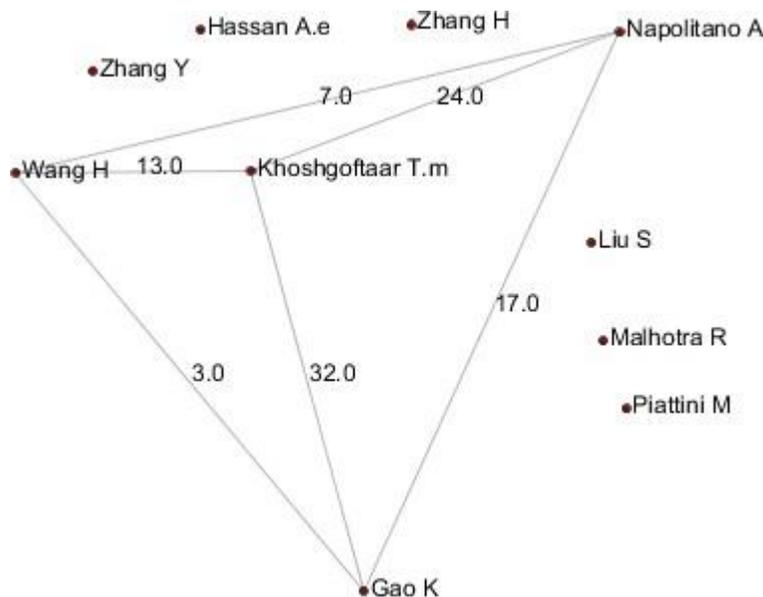
6.2 Análisis de segundo orden

En este apartado se realiza un análisis exploratorio para identificar el grado de conexión entre los elementos analizados anteriormente, esto da como resultado las redes de colaboración de: los autores, las afiliaciones y los términos claves en marco de la calidad del software.

6.2.1 Top 10 red de concurrencia de autores

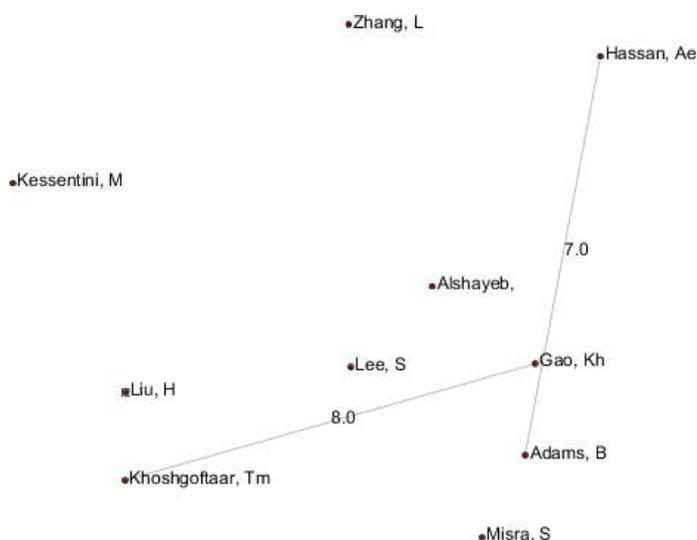
En este análisis se presenta la red de concurrencia del top 10 de autores. Los resultados permiten observar la red de colaboraciones existentes para crear publicaciones científicas entre estos autores, además, de la cantidad de documentos que han trabajado en conjunto. En la Ilustración 21 se observa la red de colaboración entre los autores de la base de datos Scopus, en donde se evidencia que 4 de los 10 autores que comprenden el top han realizado trabajos en conjunto los cuales son Napolitano A, Wang H, Gao K y Khoshgoftaar T.m. Mientras que en la Ilustración 22 se muestra la red de colaboración en ISI Web donde se evidencia que el autor Gao Kh es el autor con más colaboraciones, seguidos por los autores Khoshgoftaar T.m, Adamas B, Hassan A, los cuales comparten la misma cantidad de relaciones para la elaboración de trabajos.

Ilustración 21 Top 10 red de concurrencia de autores (Scopus)



Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

Ilustración 22 Top 10 red de concurrencia de autores (ISI Web)



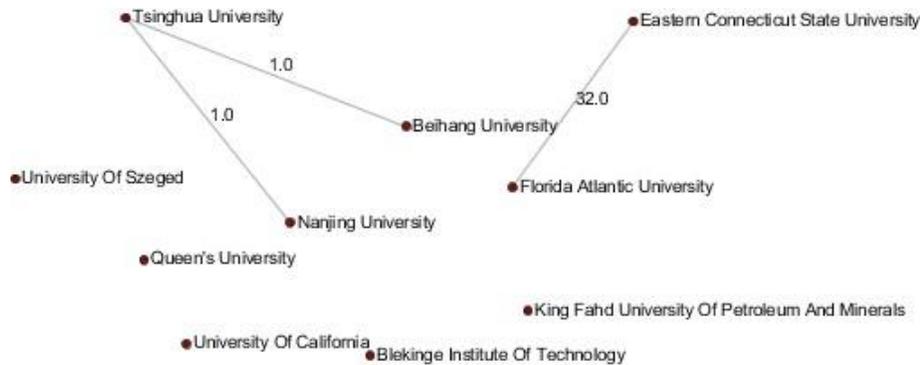
Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

6.2.2 Top 10 red de concurrencia de afiliaciones(Scopus)

En este análisis se presenta la red de concurrencia del top 10 de afiliaciones. Los resultados permiten observar las colaboraciones existentes para crear publicaciones científicas entre las afiliaciones que comprenden el top 10 (Ilustración 6). En la Ilustración 23 se observar que la universidad con mayor número de trabajos en conjunto con otras universidades es Tsinghua University con un total de 2 relaciones, seguida por Nanjing University, Beihang University, Florida Atlantic University y Eastern Connecticut State University las cuales tienen el mismo número de relaciones para la elaboración de trabajos.

Este análisis no se pudo realizar con los resultados obtenidos en la base de datos ISI Web, ya que la misma no proporciona estos datos al momento de realizar la exportación

Ilustración 23 Top 10 red de concurrencia de afiliaciones (Scopus)



Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

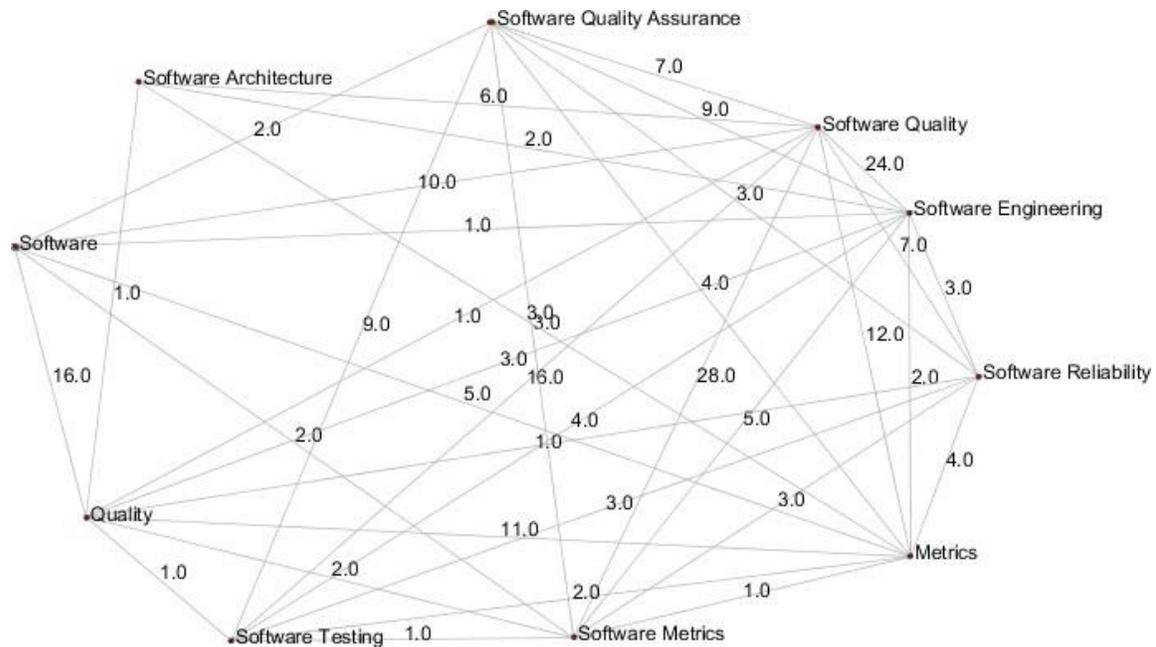
6.2.3 Top 10 red de concurrencia de palabras claves(Scopus)

En este análisis se puede observar la red de relación que existe entre los distintos términos claves utilizados en los documentos extraídos de la base de datos Scopus que más se han utilizado en las publicaciones científicas, además los cuales están relacionados con la calidad del software.

En la Ilustración 24 se presenta las distintas conexiones existentes entre los diferentes términos claves utilizados en los documentos extraídos de Scopus, en donde se observa que términos han estado trabajando juntos, con respecto al tema de Calidad de software, se evidencia que las métricas, la arquitectura y el testing son temas claves para asegurar la calidad de un software.

Este análisis no se realizó para los resultados de la base de datos ISI Web dado que, al momento de realizar la exportación de datos, estos no fueron proporcionados por la base.

Ilustración 24 Top 10 red de concurrencia de palabras claves (Scopus)

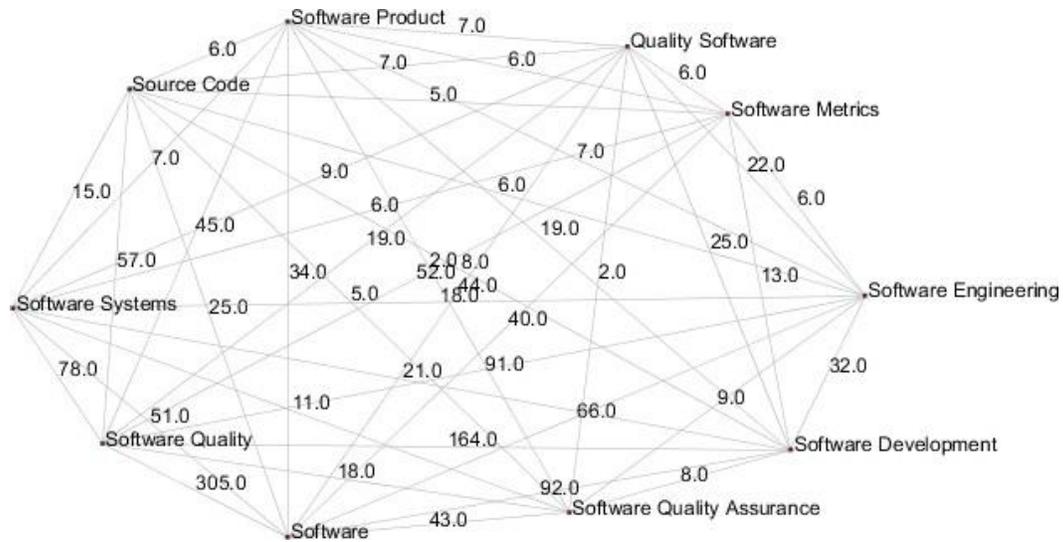


Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

6.2.4 Top 10 red de concurrencia de la bolsa de palabras

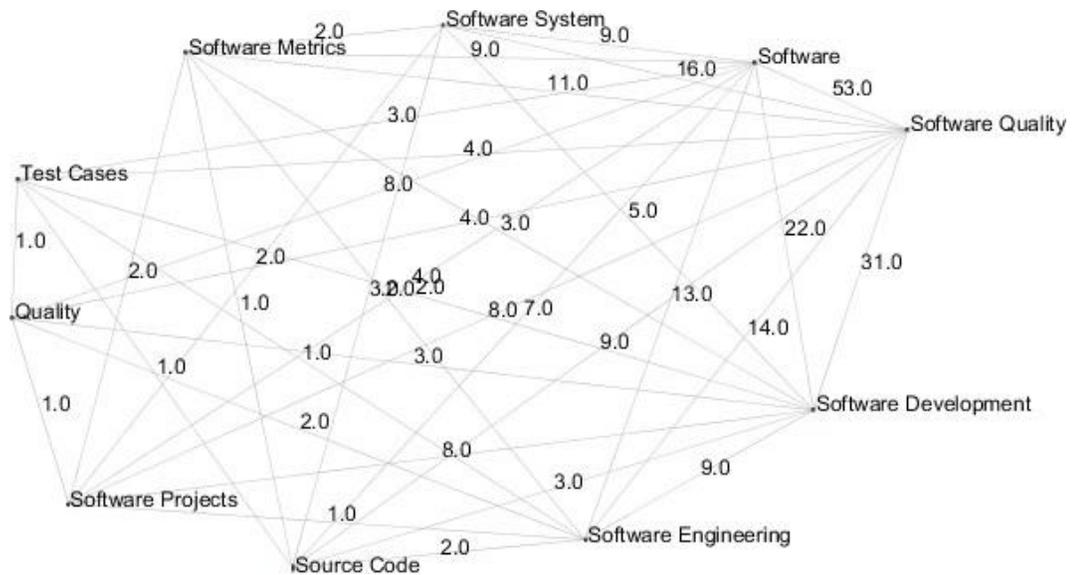
Finalmente se observa en la Ilustración 25 e Ilustración 26 los términos más usados que hacen parte de la bolsa de palabras, los cuales tienden hacer usados como referentes en los documentos o publicaciones científicas. Se observa que los temas claves que han trabajado y aplicado en conjunto, con el tema de la Calidad del software, son las métricas, la ingeniería del software y el desarrollo del software, tanto para la base datos Scopus e ISI Web.

Ilustración 25 Top 10 red de concurrencia de la bolsa de palabras (Scopus)



Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.

Ilustración 26 Top 10 red de concurrencia de la bolsa de palabras (ISI Web)



Fuente: elaboración propia con base en los resultados obtenidos en la ecuación de búsqueda.



7. Discusión

Esta pasantía tuvo como propósito construir un perfil de investigación en el área de conocimiento de la Ingeniería de Software denominada Calidad del software, con el fin de servir como guía para futuras investigaciones a través de la identificación de tendencias con respecto a la producción científica referente a la temática explorada.

De los resultados obtenidos en esta investigación se puede deducir que el aumento en las indagaciones frente a la problemática de la calidad del software se va fortaleciendo a medida que los usos de los sistemas de información se van integrando a la vida de las personas, dado que el software de alta calidad proporciona beneficios a la organización que lo produce y a la comunidad de usuarios finales esto se puede ver reflejado en la progresiva cantidad de documentos con respecto al tiempo.

Así mismo la producción de documentos científicos por afiliaciones está liderada por las universidades Florida Atlantic University, Eastern Connecticut State University, Queen's University.

De Florida Atlantic University podemos indagar que esta universidad está clasificada por la Fundación Carnegie como una universidad de investigación con alto índice de estudios, junto a las universidades Eastern Connecticut State University, Queen's University forjan innovaciones y promoviendo ciencia y tecnología en todas las disciplinas, incrementado los fondos para investigación, ampliado su infraestructura y establecido importantes vínculos con universidades líderes en investigación, esto se puede evidenciar en la Ilustración 23 donde Eastern Connecticut State University y Florida Atlantic University comparten un número de artículos de investigación elaborados en conjunto.

Otro punto significativo fue el identificar el top 10 de autores con mayor número de producción científica y su red de colaboración entre ellos, autores como Khoshgoftaar perteneciente a la Florida Atlantic University, A Hassan perteneciente a la Queen's University y Gao K perteneciente a la Eastern Connecticut State University quienes han sido muy relevantes por sus número de publicaciones de artículos científicos y que estos a su vez trabajan en conjunto con autores de otras afiliaciones (ver Ilustración 23) esto podría ser el causante de que esas universidades mantengan el tan elevada número de publicaciones que se les asigna.

Por otra parte, la producción de documentos científicos se encuentra con mayor frecuencia en los países como India y Estados Unidos, esto puede ocasionarse debido a que en la India el 72% de la producción se dedica a la exportación de software y servicios informáticos, lo cual hace invertir en la innovación e investigación en este sector; Esto, se puede ver reflejado en que en la India se encuentran 42 de las 52 empresas de software certificadas con el nivel superior en calidad. En el caso de Estados Unidos puede darse a que exporta un alto número de productos de alta tecnología a

países menos desarrollados y estos a su vez vuelven más competitivos, por ello las empresas estadounidenses se ven obligadas a invertir ciencia tecnología e investigación lo que hace que a nivel general como nación lo convierta como uno de los países con mayor potencial científico – tecnológico.

En los documentos más citados obtenidos de esta investigación se comprobó que los artículos científicos más notables son pertinentes a asuntos donde se desarrollaron sistemas de información que implican, durante el ciclo de vida de desarrollo de éstos, técnicas, normas y métodos los cuales permiten la evaluación de calidad de un producto de software y así detectar las debilidades del mismo para alimentar un proceso de mejora continua.

8. Conclusiones

De la investigación presentada podemos argumentar en primer lugar que la tendencia científica del área “Calidad del software”, esta direccionada a la obtención de métricas que ayuden a obtener una medida clara del estado actual del software en desarrollo, de igual forma temas como la refactorización de código, el código reusable y la experiencia de los usuarios finales hacen parte de los temas más investigados por las instituciones, esto se puede corroborar con los análisis de primer orden llamados Documentos más citados y Palabras claves más utilizadas por los autores.

Por otra parte, con base a los análisis descritos anteriormente se puede apreciar que el área de estudio “calidad del software” se encuentra en constante crecimiento. Haciendo un sustento en el informe presentado por SOGETI y HP [23] donde se evalúa el impacto del producto software entorno a los negocios, se aprecia el esfuerzo que hacen las empresas actualmente por asegurar la calidad de sus nuevos desarrollos software, invirtiendo en la obtención de conocimiento, herramientas, tecnologías y personal humano capacitado que hagan frente a esta problemática.

Podemos inferir que el resultado logrado fue el estudio e identificación de tendencias científicas en al área de la calidad del software, lo cual permitió la construcción de un perfil de investigación el cual tiene como objetivo ser utilizados como ejemplo o línea base para la realización de otros estudios de tendencias que se deban realizar en el marco del proyecto de investigación con el que se está construyendo el Plan Prospectivo para el desarrollo de Ciencia, Tecnología e Innovación de la universidad del magdalena. Asimismo, permite el apropiarse de las prácticas utilizadas para asegurar la calidad del producto software y poder ser implementadas para los desarrollos actuales y futuros de software realizados por la universidad de la magdalena.

En el desarrollo de esta investigación se tuvieron una serie de inconvenientes como lo fue la limpieza de los datos, debido a que la herramienta de análisis utilizada Sci2 tool no realizaba la limpieza de una manera adecuada, entonces se optó por utilizar una

herramienta llamada HAB. Pero este en sus resultados no daba los campos necesarios para realizar los análisis requeridos, de modo que se tuvo que realizar un pequeño software que permitiera realizar cruces entre los datos originales y los resultados provenientes del HAB, para de esta forma tener los campos necesarios para realizar los análisis que se presentan en los resultados. Este software fue llamado BRCC; Otro inconveniente fue debido a que la base de datos ISI Web of Knowledge (ISI Web) presenta una limitación en cuanto a los datos que permite exportar, ya que esta no exporta todos los campos necesarios para realizar los análisis correspondientes y hacer de esta manera comparaciones entre las bases de datos utilizadas.

9. Agradecimientos

El autor expresa su agradecimiento al Grupo de Investigación y Desarrollo en Tecnologías de la Información y Organizaciones de la Universidad del Magdalena; al director, Ing. ERNESTO AMARU GALVIS LISTA, Msc, PhD, por el apoyo brindado a través de esta pasantía; así mismo, al Centro de Investigación y Desarrollo de Software, por el espacio brindado para llevar a cabo la realización y ejecución del trabajo de investigación a la Ing. Mayda Patricia González Zabala, PhD, por el apoyo científico para hacer del trabajo un resultado de alta calidad así mismo a el estudiante Rafael Espitia por la colaboración brindada durante el desarrollo del proyecto.

10. Bibliografía

- [1] «concejo nacional de acreditacion,» [En línea]. Available: www.cna.gov.co/1741/articles-311056_ColombiaConstruyeSiembraFuturo.pdf. [Último acceso: 11 05 2016].
- [2] «Unimag,» [En línea]. Available: http://acreditacion.unimagdalena.edu.co/index.php/exposiciones/doc_download/286-plan-sectorial-de-ciencia-tecnologia-e-innovacion.. [Último acceso: 11 05 2016].
- [3] L. Codina, «Scopus: el mayor navegador científico de la web,» 2005. [En línea]. Available: <http://www.elprofesionaldelainformacion.com/contenidos/2005/enero/7.pdf>. [Último acceso: 11 05 2016].
- [4] UM, «Universidad de murcia,» [En línea]. Available: https://www.um.es/c/document_library/get_file?uuid=3a60c4a6-6d79-4330-a98d-9476755968a1&groupId=793464. [Último acceso: 11 05 2016].
- [5] IEEE Computer Society, SWEBOK, PIERRE BOURQUE.



- [6] R. S. Pressman, «Ingeniería del software un enfoque practico.,» Mc Graw Hill, p. 11.
- [7] «I. Sommerville. Ingeniería del software.,» United Kingdom, Pearson addison wesley, 2005, p. 6.
- [8] «C. Gil, La calidad del software como eje tematico de investigacion.,» Bogota, 2006., p. 56.
- [9] C. Sainero, «Multidoc 3.0,» [En línea]. Available:
<http://www.multidoc.es/SiteAssets/migracion/Journal/pendientedemigracion.ucm.es/info/multidoc/publicaciones/journal/pdf/bibliometria-esp.pdf>. [Último acceso: 20 06 2017].
- [1] A. L. P. y. S. W. CUNNINGHAM., «TECH MINING EXPLOITING NEW TECHNOLOGIES FOR
0] COMPETITIVE ADVANTAGE.,» Hoboken, New Jersey, John Wiley & Sons, Inc, 2005, p. 67.
- [1] U. D. MAGDALENA, « Herramienta de Analisis Bibliometrico». Colombia 2016.
1]
- [1] NIH, «national institutes of health,» [En línea]. Available:
2] <https://www.nihlibrary.nih.gov/resources/tools/science-science-sci2-tool>. [Último acceso: 21 06
2016].
- [1] S. S. Georg von Krogh, «Carrots and rainbows: Motivation and social practice in open source
3] software development,» Zurich, 2012.
- [1] N. C. Nagappan, «Don't touch my code! Examining the effects of ownership on software quality,» de
4] *19th ACM SIGSOFT Symposium on Foundations of Software Engineering*, 2011.
- [1] M. R. Heričko, «Software fault prediction metrics: A systematic literature,» 2013.
5]
- [1] S. M. Zelzer, «The medical imaging interaction toolkit: Challenges and advances: 10 years of open-
6] source development,» *International Journal of Computer Assisted Radiology and Surgery*, vol. 8, pp.
607-620, 2013.
- [1] K. T. Gao, «Choosing software metrics for defect prediction: An investigation on feature selection
7] techniques,» *Software - Practice and Experience*, vol. 41, pp. 579-606, 2011.
- [1] L. Carreno, «Analysis of user comments: An approach for software requirements evolution,» de
8] *International Conference on Software Engineering*, 2013.
- [1] F. Rahman, «Ownership, experience and defects: A fine-grained study of authorship,» de
9] *International Conference on Software Engineering*, 2011.
- [2] F. D. Rahman, «Recalling the "imprecision" of cross-project defect prediction,» de *International
0] Symposium on the Foundations of Software Engineering.*, 2012.

- [2 M. T. Kim, «A field study of refactoring challenges and benefits,» de *ACM SIGSOFT 20th International Symposium on the Foundations of Software Engineering*,, 2012.
- [2 S. M. Khan, «Factors influencing clients in the selection of offshore software outsourcing vendors: 2] An exploratory study using a systematic literature,» *Journal of Systems and Software*, vol. 84, pp. 686 - 699, 2011.
- [2 C. y. H. Sogeti, «World Quality Report,» 2016.
- 3]
- [2 A. BUSTAMANTE; E. GALVIS; , L.C. GÓMEZ, «Perfil de la investigación sobre inteligencia de 4] negocios,» *REVISTA DE LA FACULTAD DE INGENIERÍAS FÍSICO MECÁNICA (UIS Ingenierías)*, 2014.

11. Anexos

Anexo 1 Recopilación: Resumen de la metodología

Fases	Pasos	Objetivo
Inteligencia	1. Identificación del tema	Definir el área, campo o concepto a analizar y la frontera temporal y geográfica. Así como las preguntas guías del proceso.
	2. Selección de las fuentes de Información	Hacer explícitas las fuentes de información que se utilizaran para recuperar los metadatos y datos a analizar.
	3. Refinamiento de la búsqueda y recuperación de la información	Establecer la(s) ecuación(es) a utilizar para recuperar la información. Con base en la revisión de los resultados se entra en un proceso iterativo de calibración de la ecuación con base en los resultados.
Análisis y Diseño	4. Limpieza de los datos	Eliminar la redundancia y variaciones innecesarias en los datos. Se depuran los metadatos y datos dejando estos en un formato apropiado para ser utilizado por la herramienta de análisis.
	5. Análisis Básico	Empezar el análisis de los datos a través de un análisis exploratorio, la limpieza de datos (adicional al paso dos y como resultado del análisis exploratorio), el análisis de primer orden (producir listas); análisis de segundo orden (producir matrices).
	6. Análisis Avanzado	Analizar el grado de conexión entre los diferentes elementos.
Selección	7. Representación	Escoger la forma de presentar los datos dependiendo de la población objetivo del estudio.
	8. Interpretación	Hacer inferencias con base en los resultados obtenidos.

Fuente: Extraído de [24]