

## Editorial

# Una Guía Corta para Escribir Revisiones Sistemáticas de Literatura Parte 4

Esta es la última editorial de la serie dedicada al proceso de Revisión Sistemática de Literatura. En ella se aborda el proceso de documentación de la búsqueda, la selección de estudios, el análisis de los documentos seleccionados, el análisis de calidad la extracción de información y, finalmente, la síntesis de datos que corresponde a la respuesta de las preguntas de investigación a la luz de la información recopilada. Al leer estas cuatro editoriales como un todo, espero que los autores y lectores de la revista DYNALOG obtengan una visión global de la metodología que facilite el proceso de revisión de literatura y su posterior publicación.

## 1 Documentación del Proceso de Búsqueda

Una de las principales características de la revisión sistemática es que el proceso de búsqueda es repetible y sus resultados son auditables. Por ello, es requerido que se documente el proceso de búsqueda reportando la siguiente información en el documento final [1][2]:

- Bases bibliográficas, libros, bibliografías y contactos usados.
- Palabras clave y cadenas de búsqueda.
- Periodo cubierto en la revisión.

Hay diferentes opciones disponibles al seleccionar la librería digital dependiendo del área de conocimiento y es por ello que la selección debe realizarse con sumo cuidado; por ejemplo, en el caso de energía, electrónica y comunicaciones resulta obligatorio consultar la librería digital IEEE Xplore; no obstante, otras selecciones comunes incluyen: ScienceDirect, SpringerLink, Wiley online library, Emerald, ACM, Kluwer, JSTOR e ISI Web of Science. En adición, Google Scholar, Pubmed, Ebsco, Academic Search Premier, World Scientific Net, Scopus, Compendex e Inspec también pueden ser de utilidad, dependiendo del tema particular que se este abordando.

En algunas ocasiones los investigadores restringen la búsqueda a las revistas más representativas del área, tal como en [3]; sin embargo, debe recalarse que no deben olvidarse las bibliotecas tradicionales localizadas en universidades y centros de investigación, ya que ellas pueden contener documentos muy importantes para el proceso de revisión.

Otro aspecto fundamental en la búsqueda de documentos

es la selección adecuada de palabras clave, términos alternativos y posibles sinónimos [4], ya que de estos depende cuales documentos serán recuperados automáticamente. En algunos casos, cuando el Inglés no es la lengua materna, la selección de términos en un área nueva para el investigador puede ser especialmente difícil.

Otro aspecto importante es el dominio de la herramienta de búsqueda, y particularmente como estructurar la cadena de búsqueda para poder obtener el plural de los términos, considerar palabras cercanas, usar comodines y poder buscar frases exactas.

De acuerdo con [5], dos métricas pueden ser usadas para analizar la estrategia automática de búsqueda: la precisión, definida como el número de estudios relevantes recuperados sobre el total de estudios recuperados; y la sensibilidad, que es el número de estudios relevantes recuperados sobre el total de estudios relevantes existentes. Una estrategia óptima de búsqueda tiene usualmente una precisión entre el 20% y el 30% y una sensibilidad entre el 80% y el 99% [5]. Esto significa que para un total de cien estudios finalmente seleccionados, el investigador tuvo que revisar, al menos, entre 400 y 500 estudios recuperados automáticamente; esto significa también, que cuando la búsqueda no esta bien estructurada, se deben analizar mucho más de 500 estudios para seleccionar los mismos 100 estudios.

Una de las herramientas preferidas para realizar la búsqueda es Scopus ([www.scopus.com](http://www.scopus.com)) porque incluye más de 219.000 títulos de aproximadamente 5.000 editoriales científicos, incluyendo las mayores editoriales en el mundo, tales como IEEE, Elsevier, Emerald, Springer y Wiley; adicionalmente, Scopus tiene un importante conjunto de herramientas para el análisis de literatura. Algunos consejos para crear búsquedas útiles en Scopus son los siguientes:

- Se pueden usar los operadores booleanos AND, OR y AND NOT en la cadena de búsqueda.
- La expresión  $w_1 W/n w_2$  indica que los términos  $w_1$  y  $w_2$  no pueden estar separados por más de  $n$  términos.
- La expresión  $w_1 PRE/n w_2$  especifica que el término  $w_1$  precede el término  $w_2$  por no más de  $n$  términos.
- El asterisco es un comodín; por ejemplo, la cadena `forecast*` encuentra las palabras `forecasts`, `forecasting` y `forecasted`.

- Sólo es necesario utilizar el singular de una palabra ya que el plural y el posesivo son automáticamente considerados.
- Use las llaves (“{” y “}”) para indicar frases exactas.
- Use las comillas dobles para indicar búsqueda aproximada. En este caso se puede usar el “\*” como comodín.

En Scopus también es posible ejecutar varias búsquedas y luego combinarlas usando los operadores lógicos OR, AND y AND NOT. Adicionalmente, Scopus presenta opciones para excluir o limitar la búsqueda por autores, años, instituciones, áreas temáticas, tipos de documento, palabras clave, afiliaciones, etc. También es posible exportar información bibliográfica a administradores de referencias. Recientemente, Elsevier incorporó el administrador bibliográfico Mendeley a ScienceDirect tal que es posible importar directamente a Mendeley documentos en formato pdf e información bibliográfica.

En el caso de la librería digital IEEE Xplorer, algunos consejos para realizar búsquedas avanzadas son los siguientes:

- Se pueden usar los operadores booleanos AND, OR y NOT en la cadena de búsqueda.
- Los signos de puntuación son ignorados en las cadenas de búsqueda.
- Sólo es necesario utilizar el singular de una palabra ya que el plural y el posesivo son automáticamente considerados. También se consideran automáticamente las palabras en inglés americano e inglés británico.
- Las comillas dobles se usan para indicar frases exactas.
- El “\*” es usado como comodín.

Frecuentemente, el uso de librerías digitales y bases de datos no es suficiente para localizar la literatura más relevante y se hace necesario consultar en las librerías de las universidades y centros de investigación en busca de libros, reportes de investigación, tesis de maestría, disertaciones de doctorado y cualquier otro tipo de literatura que pueda ser útil en su investigación. Se deben utilizar las referencias relevantes ya localizadas para poder extender la lista de palabras clave y refinar la cadena de búsqueda.

Cuando el reporte es escrito, es común presentar la cadena de búsqueda que fue usada en la librerías digitales con el fin de hacer la búsqueda reproducible, verificable y explícita y para poder actualizarla después [4]. Por ejemplo, en [6] se reporta el proceso de búsqueda de la siguiente forma:

Our five research questions contain the following key words:

“Software Engineer, Software Engineering, Motivation, De-Motivation, Productive, Characteristics, outcome, Model”. A list of synonyms was constructed for each of these words, as in the example for research question 1 which contains keywords ‘Software Engineer’ and ‘Characteristics’:

keywords((engineer\* OR developer\* OR professional\* OR programmer\* OR personnel OR people OR analyst\* OR team leader\* OR project manager\* OR practitioner\* OR maintainer\* OR designer\* OR coder\* OR tester\*) AND characteristic\* OR types OR personality OR human factors OR different OR difference\* OR psychology OR psychological factors OR motivator\* OR prefer\* OR behavio\*r\*)

## 2 Selección de los Estudios

Es el proceso de selección de los estudios que serán utilizados finalmente en la revisión. Cuando la búsqueda es ejecutada usando las cadenas de búsquedas diseñadas en el paso anterior, una gran cantidad de documentos es recuperada automáticamente, pero únicamente unos pocos documentos serán relevantes para la revisión. Una búsqueda bien diseñada permite recuperar la literatura más relevante relacionada con el área de investigación; pero también muchos documentos irrelevantes o por fuera del foco del estudio; de esta forma, es obligatorio el realizar una depuración manual de los documentos recuperados automáticamente.

De esta forma, cuando el protocolo de investigación es escrito, se deben definir dos listas en la sección donde se describe el protocolo de selección, una para los criterios de inclusión y otra para los criterios de exclusión [1][2]. La claridad es obligatoria en la definición de los criterios para que el proceso pueda ser auditado o reproducido [7]. Los documentos son comúnmente descartados mediante la lectura de sus títulos, resúmenes o conclusiones [7]; pero en caso de duda, se requiere una lectura más profunda. Es muy común que se realice la exclusión basándose en el idioma del documento o cuando los documentos son opiniones, puntos de vista o anécdotas [6]. El filtrado manual de los documentos es realizado por la aplicación de los criterios de inclusión y exclusión con el fin de poder determinar cuales serán los documentos finalmente usados en la revisión [1][2]. Cuando hay documentos traslapados en su contenido o duplicados, se debe seleccionar únicamente el más completo de ellos y el resto de documentos debe ser descartado; véase por ejemplo [8]. Los directores de tesis, compañeros, expertos y contactos son una fuente muy valiosa para determinar si se han omitido documentos importantes [7].

Otro aspecto muy importante es el uso apropiado de herramientas de software para el manejo de las referencias [9][4], tales como Zotero, Mendeley, Papers, Sente, Qiqqa, EndNote, Reference Manaer o ProCite, ya que el número final de documentos puede ser muy grande. Algunos gestores de referencias bibliográficas como Papers o Mendeley, permiten que el usuario pueda administrar los documentos, agregar notas y resaltar o subrayar textos directamente en los documentos.

### 3 Análisis de los Documentos Seleccionados

En la sección de resultados de los informes de SLR, se describe el proceso de recuperación de documentos y se analizan los documentos finalmente seleccionados como un todo; en esta sección del reporte y para el caso de las revisiones sistemáticas de literatura, muchas de las preguntas de investigación comúnmente usadas en los estudios de mapeo sistemático son indirectamente respondidas; esto es, las preguntas como tal no son formuladas pero si se reporta el resultado de su análisis. En la Tabla I se describen los tipos comunes de tablas usada para analizar los documentos, así como también las columnas típicas de dichas tablas.

Para mediar la contribución de un autor, es posible calcular el puntaje  $S$  que depende del número de autores del estudio ( $n$ ) y de la posición ordinal ( $i$ ) del nombre del autor en la lista de autores del documento analizado, tal como se muestra en [10]:

$$S = \frac{1.5^{n-1}}{\sum_{i=1}^n 1.5^{n-i}}$$

Tenga en cuenta que un puntaje de 1,0 es obtenido cuando el documento tiene un solo autor; se obtienen puntajes de 0,60 (primer autor) y de 0,40 (segundo autor) cuando el documento tiene dos autores, y así sucesivamente. También se pueden usar otros indicadores bibliométricos para analizar la literatura, tal como el factor de impacto o el índice de colaboración.

La mayoría de artículos reportando resultados de SLR, indican el número de artículos recuperados automáticamente, seleccionados usando los criterios de inclusión, descartados usando los criterios de exclusión, y finalmente seleccionados usando el análisis de calidad; véase por ejemplo [11].

### 4 Realización del Análisis de Calidad

Los resultados del análisis de calidad de los estudios seleccionados pueden ser usados como [1][2]:

- Un indicador de la calidad para determinar cómo ha progresado el campo de investigación; este es un indicador de la madurez de la investigación.
- Un criterio para determinar que estudios deben ser excluidos cuando no cumplen con un nivel mínimo de calidad.
- Un criterio para medir la importancia que debe tener cada estudio en la respuesta de las preguntas de investigación.
- Un mecanismo para explicar la diferencia entre los resultados de estudios similares.
- Un criterio para sugerir nueva investigación.

Tabla 1.

Tablas comúnmente usadas para analizar los documentos seleccionados en revisiones sistemáticas de literatura y mapeos sistemáticos

Tabla	Columnas Típicas
Número de documentos seleccionados por librería digital	Librería Digital Número Total de Documentos Número de Documentos Incluidos Numero de Documentos Excluidos
Estudios Seleccionados	Autores Librería Digital Título Año Fuente Número de Citaciones Tipo de Documento Tipo de Estudio Tema del Estudio Objetivo del Estudio Método de Investigación [Otras columnas... ]
Número de Artículos Publicados por Revista	Nombre de la Revista Numero de Artículos Porcentaje (%)
Número de Artículos Publicados por Editorial	Editorial Número de Artículos Porcentaje (%)
Artículos más Citados	Autores Año Fuente Número de Citaciones
Número de Artículos por Autor	Autor Filiación Número de Citas por Autor Numero de Artículos Puntaje $S$
Número de Artículos por País	País Número de Filiaciones Diferentes Número de Autores Número de Artículos Porcentaje de Artículos (%) Puntaje Total $S$ por País
Número de artículos por Continente	Continente Número de Filiaciones Diferentes Número de Autores Número de Artículos Porcentaje de Artículos (%) Puntaje Total $S$ por Continente
Número de Artículos Publicados por Año	Año Número de Artículos Porcentaje de Artículos (%) Porcentaje Acumulado (%)

Nota: Algunas de estas tablas pueden ser presentadas alternativamente como gráficas.

El análisis de calidad es dependiente de los objetivos de la revisión seleccionada. Las Tablas 5 y 6 de la ref. [1] contienen una recopilación de preguntas usadas en el análisis de calidad de diferentes estudios.

Una forma de realizar el análisis de la calidad es considerar en que proporción cada estudio seleccionado

permite responder las preguntas de investigación [3]; comúnmente se usa una escala nominal: Y (sí), N (no) y P (parcialmente). Para obtener un puntaje final para cada estudio, se asignan valores a la escala nominal, por ejemplo, se asigna 1,0 a Y, 0,5 a P y 0,0 a N, y luego se suman los puntos asignados a cada estudio. Los resultados de la evaluación de la calidad son comúnmente presentados en forma de una tabla detallada.

## 5 Extracción de los Datos

La información para responder las preguntas de investigación y realizar la evaluación de la calidad es recolectada en forma de tablas o cualquier otro tipo de sistema de toma de notas [1][7]. Las tablas son diseñadas en una manera consistente con las preguntas de investigación y permiten enfatizar las similitudes y diferencias entre los estudios [1].

Cuando un gestor de referencias como Mendeley o Paper es usado, es posible resaltar texto con diferentes colores, subrayar y agregar anotaciones y comentarios en los documentos con el fin de facilitar la extracción de datos y su posterior revisión. Una selección común es usar diferentes colores para marcar las partes donde cada respuesta a una pregunta de investigación específica aparece en el documento.

## 6 Síntesis de los Datos

En este paso [2], las preguntas de investigación son respondidas usando la información recolectada. Cuando los estudios primarios son cuantitativos, la síntesis puede ser realizada usando técnicas estadísticas, tales como el meta-análisis; pero cuando los estudios son cualitativos, la síntesis es narrativa (descriptiva) resumiendo los principales hechos. En la síntesis descriptiva, es usual preparar tablas y gráficos que resuman la información relevante respecto a la pregunta o preguntas de investigación consideradas. Dos aproximaciones pueden ser usadas para responder estudios cualitativos [12]: centradas en los conceptos o centradas en los autores.

## Referencias

- [1] Kitchenham, B.A. and Charters, S., Guidelines for performing systematic literature reviews in software engineering. Technical Report EBSE-2007-01, 2007.
- [2] S. Sorrell, Improving the evidence base for energy policy: The role of systematic reviews, *Energy Policy*, 35 (3), pp. 1858-1871, 2007. <http://dx.doi.org/10.1016/j.enpol.2006.06.008>
- [3] Kitchenham, B., Brereton, O.P., Budgen, D., Turner, M., Bailey, J. and Linkman, S., Systematic literature reviews in software engineering – A systematic literature review, *Information and Software Technology*, 51 (1), pp. 7-15, 2009. <http://dx.doi.org/10.1016/j.infsof.2008.09.009>
- [4] Bolderston, A., Writing an Effective Literature Review, *Journal of Medical Imaging and Radiation Sciences*, 39 (2), pp. 86-92, 2008. <http://dx.doi.org/10.1016/j.jmir.2008.04.009>
- [5] Zhang, H., Babar, M.A. and Tell, P., Identifying relevant studies in software engineering, *Information and Software Technology*, 53 (6), pp. 625-637, 2011. <http://dx.doi.org/10.1016/j.infsof.2010.12.010>
- [6] Beecham, S., Baddoo, N., Hall, T., Robinson, H. and Sharp, H., Motivation in software engineering: A systematic literature review, *Information and Software Technology*, 50 (9-10), pp. 860-878, 2008. <http://dx.doi.org/10.1016/j.infsof.2007.09.004>
- [7] Randolph, J.J., A guide to writing the dissertation literature review, *Practical Assessment, Research & Evaluation*, 4 (13), pp. 1-13, 2009.
- [8] Khurum, M. and Gorschek, T., A systematic review of domain analysis solutions for product lines, *The Journal of Systems and Software*, 82 (12) pp. 1982-2003, 2009. <http://dx.doi.org/10.1016/j.jss.2009.06.048>
- [9] Pautasso, M., Ten simple rules for writing a literature review, *PLoS Computational Biology*, 9 (7), pp. 1-4, 2013. <http://dx.doi.org/10.1371/journal.pcbi.1003149>
- [10] Yuan, H. and Shen, L., Trend of the research on construction and demolition waste management, *Waste Management*, 31 (4), pp. 670-679, 2011. <http://dx.doi.org/10.1016/j.wasman.2010.10.030>
- [11] Hauge, Ø., Ayala, C. and Conradi, R., Adoption of open source software in software-intensive organizations – A systematic literature review, *Information and Software Technology*, 52 (11), pp. 1133-1154, 2010. <http://dx.doi.org/10.1016/j.infsof.2010.05.008>
- [12] Webster, J. and Watson, R.T., Editorial. Analyzing the past to prepare for the future: Writing a literature review, *MIS Quarterly*, 26 (2), pp. xiii-xxiii, 2002.

Juan D. Velásquez, MSc, PhD  
 Profesor Titular  
 Universidad Nacional de Colombia  
 E-mail: jdvelasq@unal.edu.co  
<http://orcid.org/0000-0003-3043-3037>