

**PEDECIBA Informática**  
**Instituto de Computación – Facultad de Ingeniería**  
**Universidad de la República**  
**Montevideo, Uruguay**

---

## **Reporte Técnico RT 15-08**

---

**CONCEPTOS DE INGENIERIA DE  
SOFTWARE BASADO EN EVIDENCIAS**

**SEBASTIAN PIZARD**

**FERNANDO ACERENZA**

**VANESSA CASELLA**

**SILVANA MORENO**

**DIEGO VALLESPER**

2015

Sebastian Pizard, Fermamdo Aceremza, Vanessa Casella,  
Silvana Moreno, Diego Vallespir  
ISSN 0797-6410  
Tesis de Maestría en Informática  
**Reporte Técnico RT 1508**  
PEDECIBA  
Instituto de Computación – Facultad de Ingeniería  
Universidad de la República.  
Montevideo, Uruguay, 2015



# Conceptos de Ingeniería de Software Basada en Evidencias

*Versión 1.0*

Sebastián Pizard  
Fernando Acerenza  
Vanessa Casella  
Silvana Moreno  
Diego Vallespir

Reporte Técnico InCo/Pedeciba-2015 TR:15-08  
Setiembre, 2015

## **Resumen**

En este documento se presenta la definición de la Ingeniería de Software Basada en Evidencias y se describen sus objetivos. De forma más práctica, se incluye un resumen de los métodos más utilizados para la concreción de resultados, estos son: Revisiones Sistemáticas de Literatura, Estudios de Mapeo y Revisiones Terciarias.

## Control de Versiones del Documento

<b>Versión</b>	<b># Reporte</b>	<b>Fecha</b>	<b>Cambios</b>
Versión 1	Sin número	Setiembre 2015	Versión inicial

## Índice general

1.Introducción.....	5
2.Ingeniería de Software Basada en Evidencias.....	5
3.Revisiones Sistemáticas de Literatura.....	6
3.1.Motivación.....	6
3.2.Definición y Etapas.....	6
3.3.Planificar la Revisión.....	7
3.3.1.Identificar la necesidad de una revisión.....	7
3.3.2.Puesta en marcha de la revisión.....	8
3.3.3.Especificar las preguntas de investigación.....	8
3.3.4.Desarrollar el protocolo de revisión.....	9
3.3.5.Evaluar el protocolo de revisión.....	11
3.4.Realizar la Revisión.....	11
3.4.1.Identificar la investigación.....	11
3.4.2.Seleccionar los estudios primarios.....	12
3.4.3.Evaluar la calidad de los estudios.....	13
3.4.4.Extraer datos.....	14
3.4.5.Sintetizar los datos.....	15
3.5.Informar sobre la revisión.....	16
3.5.1.Especificar los mecanismos de difusión.....	16
3.5.2.Formatar el informe principal.....	16
3.5.3.Evaluar el informe.....	18
4.Estudios de Mapeo.....	18
4.1.Diferencias entre un estudio de mapeo y una SLR.....	18
4.2.Esquemas de clasificación.....	19
4.3.Presentación de resultados.....	20
5.Revisiones Terciarias.....	22
6.Anotaciones sobre librerías digitales y motores de búsqueda.....	22
6.1.ACM Digital.....	23
6.2.IEEEExplore.....	25
6.3.ScienceDirect.....	28
6.4.Scopus.....	30
6.5.SpringerLink.....	31
7.Referencias.....	31
8.Material complementario.....	34

## 1. Introducción

Este reporte tiene como objetivo presentar los fundamentos de la Ingeniería de Software Basada en Evidencias (o ESBE por sus siglas en inglés), y en particular se incluye una breve guía para realizar revisiones sistemáticas en ingeniería de software. Se pretende que este documento sea utilizado por Proyectos de Grado de la carrera Ingeniería en Computación, Maestrías, Doctorados, etc. que se encuentren realizando trabajos de ESBE con el Grupo de Ingeniería de Software (GrlS). De esta manera los estudiantes de grado y posgrado pueden usar este documento como punto de partida para comprender la ESBE. Además, pueden incluir este documento como parte de su informe de proyecto o su Tesis evitando tener un enfoque distinto de la ESBE en cada Proyecto de Grado y/o Tesis.

En la sección 2 se presenta la definición de la ESBE así como sus motivaciones. En la sección 3 se describen las revisiones sistemáticas de literatura, en la sección 4 los estudios de mapeo. En la sección 5 se introduce a las revisiones terciarias. En la sección 6 se incluyen anotaciones sobre librerías digitales y motores de búsqueda.

Por último se incluyen 3 anexos seleccionados del sitio de ingeniería de software basado en evidencias – Evidence-Based Software Engineering (<http://community.dur.ac.uk/ebse/>). Estos corresponden a un plantilla para revisiones sistemáticas de literatura, una plantilla para estudios de mapeo y un ejemplo de una revisión sistemática de literatura. Los tres documentos no ha sido modificados y se reproducen bajo la licencia de Creative Commons (<http://creativecommons.org/licenses/by-sa/2.0/uk/>).

## 2. Ingeniería de Software Basada en Evidencias

La ingeniería de software basada en evidencias tiene como propósito mejorar la toma de decisiones relacionada al desarrollo y mantenimiento de software integrando la mejor evidencia actual de la investigación con experiencias prácticas y valores humanos. Para lograr esto, alienta un fuerte énfasis en el rigor metodológico involucrando los siguientes cinco pasos [1]:

- 1) Convertir un problema relevante o una necesidad de información en una pregunta que pueda ser respondida.
- 2) Buscar en la literatura la mejor evidencia para responder a esa pregunta.
- 3) Evaluar de forma crítica la validez, el impacto y la aplicación de la evidencia.
- 4) Integrar la evidencia evaluada con la experiencia práctica y los valores y circunstancias de los interesados.
- 5) Evaluar el desempeño y buscar maneras de mejorarlo.

La EBSE surge como una analogía a la práctica médica en donde el paradigma ha cambiado drásticamente la investigación. En esa rama de las ciencias se utilizan revisiones para resumir, por ejemplo, los resultados de una nueva droga presente en distintos medicamentos probados en diferentes lugares del mundo. La rigurosidad metodológica hace que los resultados sean más confiables ya que es posible estudiar el procedimiento llevado a cabo para su obtención así como también reproducirlo.

El mecanismo normal para la identificación y la agregación de evidencia de investigación es la Revisión Sistemática de Literatura.



### 3. Revisiones Sistemáticas de Literatura

#### 3.1. Motivación

La necesidad de agregar evidencia desde múltiples estudios empíricos tiene varias motivaciones. Seguramente al comenzar una investigación se quiera utilizar toda la investigación previa relacionada con el fin de no partir de cero. También es posible querer contemplar los resultados de varios estudios particulares en conjunto para contestar preguntas amplias que no podrían contestarse de forma individual. Este proceso de recolección y síntesis debe contar con cierto estándar científico y rigurosidad metodológica.

Las revisiones sistemáticas de literatura permiten recolectar y sintetizar evidencia de distintas fuentes. La característica clave que las distingue de las revisiones tradicionales narrativas (o clásicas) es su intento explícito de minimizar las posibilidades de llegar a conclusiones erradas, que puedan resultar de sesgo en los estudios primarios o en el proceso mismo de revisión [2].

Sobre este tema:

- Una breve descripción de las revisiones sistemáticas de literatura en ingeniería de software puede encontrarse en [3] o en el capítulo 4 de [4], la referencia más importante es la guía de Kitchenham [5].
- Shull et al. presentan una breve pero útil introducción a los distintos métodos empíricos utilizados en la investigación en ingeniería de software [6].
- Kitchenham et al. brinda una descripción más detallada de las razones e importancia de las revisiones sistemáticas, así como de sus ventajas y desventajas [5].

#### 3.2. Definición y Etapas

Una revisión sistemática de literatura (o SLR por sus siglas en inglés) es un método para identificar, evaluar e interpretar todas las investigaciones pertinentes a una determinada pregunta de investigación, área temática o fenómeno de interés. Los estudios individuales que contribuyen a una revisión sistemática son llamados estudios primarios; mientras que una revisión sistemática es por tanto un estudio secundario [5].

Una SLR tiene ciertas etapas discretas contempladas en tres fases principales [5]:

##### 1) Planificar la Revisión

- Identificar la necesidad de una revisión.
- Poner en marcha la revisión (opcional).
- Especificar las preguntas de investigación.
- Desarrollar el protocolo de revisión.
- Evaluar el protocolo de revisión (recomendada).

## 2) Realizar la Revisión

- Identificar la investigación.
- Seleccionar los estudios primarios.
- Evaluar la calidad de los estudios.
- Extraer datos y monitorear.
- Sintetizar los datos.

## 3) Informar la Revisión

- Especificar los mecanismos de difusión.
- Formatear el informe principal.
- Evaluar el informe (recomendada).

La lista de etapas no es estrictamente secuencial y algunas pueden repetirse más de una vez y pueden involucrar iteración o reimplementación.

### 3.3. Planificar la Revisión

#### 3.3.1. Identificar la necesidad de una revisión

Antes de comenzar es necesario chequear si no existen revisiones previas o en curso y si una nueva revisión es necesaria. Esta tarea se puede realizar buscando revisiones en los motores de búsqueda relacionados al área temática. Por ejemplo si se quiere realizar una revisión sobre educación de ingeniería de software se puede chequear revisiones existentes utilizando la cadena "*software engineering education systematic literature revision*" en diferentes motores de búsqueda. Si se identifican una o más revisiones que contestan la pregunta de interés, entonces se deben evaluar si cuentan con calidad suficiente.

Una evaluación crítica puede ser realizada utilizando una lista de verificación o checklist. A continuación se incluye una de ejemplo [7]:

- ¿Fue la pregunta de investigación claramente definida en términos de población, intervenciones, comparaciones, resultados y diseño de estudios?
- ¿Fue la estrategia de búsqueda adecuada y apropiada?  
¿Hubo alguna restricción de lenguaje, estado o fecha de publicación?
- ¿Se realizaron acciones preventivas para minimizar sesgo y errores en el proceso de selección de estudios?
- ¿Se usó un criterio apropiado para evaluar la calidad de los estudios primarios y se realizaron acciones preventivas para minimizar sesgo y errores en el proceso de evaluación de la calidad?
- ¿Se realizaron acciones preventivas para minimizar sesgo y errores en el proceso de extracción de datos?
- ¿Son adecuados los detalles presentados de cada estudio primario?
- ¿Son apropiados los métodos utilizados para la síntesis de datos? ¿Hay diferencias entre los estudios evaluados? ¿Se agruparon los estudios, y si fue así: era apropiado y significativo hacerlo?
- ¿Reflejan las conclusiones de los autores la evidencia que se revisó?

### 3.3.2. Puesta en marcha de la revisión

A veces una organización requiere de información sobre un tema específico pero no cuenta con el tiempo o la experiencia para realizar una revisión sistemática propia. En tales casos, podrá encargar a investigadores la revisión sistemática de literatura sobre el tema. En estos casos, la organización deberá presentar un documento de puesta en marcha especificando el trabajo requerido. Un documento de este estilo puede tener carácter contractual y debe explicitar cabalmente los requisitos de la organización.

### 3.3.3. Especificar las preguntas de investigación

La especificación de las preguntas de investigación es la parte más importante de cualquier revisión sistemática ya que éstas constituyen la base de toda la revisión: definen qué estudios serán incluidos o excluidos de la revisión, definen los datos que se deberían extraer de la literatura revisada y finalmente son respondidas en el reporte final de la revisión utilizando el proceso de síntesis de datos elegido.

De acuerdo a Kitchenham [5], es conveniente crear las preguntas correctas. En general una pregunta correcta es aquella que:

- Es significativa e importante tanto para investigadores como para profesionales.
- Dará lugar a cambios o incrementará la confianza en las prácticas actuales de la ingeniería de software.
- Identificará las discrepancias entre las creencias comúnmente aceptadas y la realidad.

Aspectos a considerar al crear las preguntas de investigación incluyen [5], [4]:

- La *población* en la cual se recolecta la evidencia. Por ejemplo el grupo de programadores que probaron un nuevo método para diseñar software.
- La *intervención* aplicada en el estudio empírico. Por ejemplo, el nuevo método para diseñar software.
- La *comparación* contra la cual se compara la intervención. Por ejemplo, el método antiguo que los programadores utilizaban para diseñar software.
- Los *resultados*. Además de ser estadísticamente satisfactorios deben ser significativos desde el punto de vista práctico. Por ejemplo, seguramente se pueda obviar en los resultados que el nuevo método de diseño requiere el uso de hojas oficio en lugar de A4.
- El *contexto* del estudio. En general es una extensión de la población, puede incluir si fue realizado en la industria o la academia, en qué segmento de la industria o por ejemplo, los incentivos otorgados a los sujetos.
- Los *diseños experimentales* a incluir en las preguntas de investigación. Quizás se requiera restringir a revisar sólo estudios primarios correspondientes a casos de estudio, por ejemplo.

#### Recomendaciones [8], [9]

- Es probable que sea necesario revisar las preguntas durante el desarrollo del protocolo, a medida que aumente el entendimiento del problema.
- Quizás realizar un estudio de mapeo previo a la revisión sirva de ayuda para determinar el alcance de las preguntas de investigación.
- Es conveniente evaluar que los motores de búsqueda sean capaces de ejecutar las cadenas de búsqueda.

#### 3.3.4. Desarrollar el protocolo de revisión

Un protocolo de revisión especifica los métodos que se utilizarán para llevar a cabo una revisión sistemática específica y su definición de antemano tiene como objetivo minimizar cualquier sesgo en el proceso. Por ejemplo, sin un protocolo, es posible que la selección y/o el análisis de los estudios individuales se vean influenciados por las expectativas del investigador.

Los componentes de un protocolo incluyen todos los elementos de la revisión más alguna información adicional de planificación:

- Antecedentes y justificación de la revisión.
- Preguntas de investigación.
- Estrategia de búsqueda para estudios primarios. Debe incluir los términos de búsqueda y recursos donde se realizará la búsqueda. Los recursos incluyen librerías digitales, revistas específicas, y actas de congresos.
- Criterios de selección de estudios. Usados para determinar cuáles estudios son incluidos o excluidos de la revisión sistemática
- Procedimientos de selección de estudios. El protocolo debe describir cómo se aplicará el criterio de selección, por ejemplo, cuántos asesores evaluarán cada estudio primario, y cómo se resolverán los desacuerdos entre los evaluadores.
- Procedimientos y listas de verificación para evaluar la calidad.
- Estrategia de extracción de datos.
- Síntesis de los datos extraídos.
- Estrategia de difusión.
- Calendario del proyecto de revisión.

En la Tabla 1 se incluye una posible estructura para el protocolo [9].

- |                                     |
|-------------------------------------|
| 1. Question Formularization         |
| 1.1. Question Focus                 |
| 1.2. Question Quality and Amplitude |
| - Problem                           |
| - Question                          |
| - Keywords and Synonyms             |
| - Intervention                      |
| - Control                           |
| - Effect                            |
| - Outcome Measure                   |
| - Population                        |
| - Application                       |
| - Experimental Design               |

<p>2. Sources Selection</p> <p>2.1. Sources Selection Criteria Definition</p> <p>2.2. Studies Languages</p> <p>2.3. Sources Identification</p> <ul style="list-style-type: none"> <li>- Sources Search Methods</li> <li>- Search String</li> <li>- Sources List</li> </ul> <p>2.4. Sources Selection after Evaluation</p> <p>2.5. References Checking</p> <p>3. Studies Selection</p> <p>3.1. Studies Definition</p> <ul style="list-style-type: none"> <li>- Studies Inclusion and Exclusion Criteria Definition</li> <li>- Studies Types Definition</li> <li>- Procedures for Studies Selection</li> </ul> <p>3.2. Selection Execution</p> <ul style="list-style-type: none"> <li>- Initial Studies Selection</li> <li>- Studies Quality Evaluation</li> <li>- Selection Review</li> </ul> <p>4. Information Extraction</p> <p>4.1. Information Inclusion and Exclusion Criteria Definition</p> <p>4.2. Data Extraction Forms</p> <p>4.3. Extraction Execution</p> <ul style="list-style-type: none"> <li>- Objective Results Extraction <ul style="list-style-type: none"> <li>i) Study Identification</li> <li>ii) Study Methodology</li> <li>iii) Study Results</li> <li>iv) Study Problems</li> </ul> </li> <li>- Subjective Results Extraction <ul style="list-style-type: none"> <li>i) Information through authors</li> <li>ii) General Impressions and Abstractions</li> </ul> </li> </ul> <p>4.4. Resolution of divergences among reviewers</p> <p>5. Results Summarization</p> <p>5.1. Results Statistical Calculus</p> <p>5.2. Results Presentation in Tables</p> <p>5.3. Sensitivity Analysis</p> <p>5.4. Plotting</p> <p>5.5. Final Comments</p> <ul style="list-style-type: none"> <li>- Number of Studies</li> <li>- Search, Selection and Extraction Bias</li> <li>- Publication Bias</li> <li>- Inter-Reviewers Variation</li> <li>- Results Application</li> <li>- Recommendations</li> </ul>
<p>Tabla 1 – Estructura común de los protocolos de revisiones sistemáticas de literatura.</p>

#### Recomendaciones [8], [3]

- Conviene que todos los miembros de la revisión sistemática participen activamente en el desarrollo del protocolo de revisión.
- Resulta esencial realizar una prueba piloto del protocolo de investigación. Esto permite encontrar errores en los procedimientos de recolección y agregación de los datos. Es posible que surjan consideraciones que hagan cambiar la metodología que se pretendía usar para contestar las preguntas de investigación.
- Es probable que el protocolo sea modificado a lo largo del proceso de la revisión, por tanto se deben documentar y justificar todos los cambios realizados.

### 3.3.5. Evaluar el protocolo de revisión

Ya que el protocolo es un elemento crítico de cualquier revisión sistemática, los investigadores deberían acordar un procedimiento para su evaluación.

Se pueden utilizar los mismos criterios utilizados para la identificación de la revisión. Además, se puede chequear la consistencia interna del protocolo para confirmar que: las cadenas de búsqueda fueron derivadas correctamente; los datos a extraer y el procedimiento de análisis de los datos responderán adecuadamente a las preguntas de investigación.

#### Recomendaciones [8]

- Para la extracción de datos conviene contar con un documento breve que contenga únicamente la definición de los datos y las guías de extracción de datos del protocolo.
- Es necesario que se acuerde un proceso de validación separado de las pruebas piloto del protocolo. Sería ideal que este proceso fuera realizado por revisores externos.

## 3.4. Realizar la Revisión

### 3.4.1. Identificar la investigación

En esta etapa se establece y utiliza una estrategia de búsqueda para obtener una lista de todas las publicaciones relevantes para las preguntas de investigación. La estrategia de búsqueda tiene que ser la definida en el protocolo de la revisión y debe estar publicada de forma transparente y replicable.

Es bueno descomponer las preguntas de investigación e identificar cadenas iniciales de búsqueda de acuerdo a población, intervención, comparación, resultados, contexto y diseño del estudio. Además conviene crear listas de sinónimos, abreviaciones y ortografías alternativas.

Otra alternativa, a veces utilizada como complemento, es el uso de las referencias en la búsqueda de nuevos estudios. Se parte de un estudio primario relevante encontrado previamente y se utilizan sus referencias para encontrar nuevos estudios relevantes (*backward snowballing*) o se buscan que estudios lo ubican en sus referencias (*forward snowballing*) [4].

Al crear la estrategia de búsqueda se debe perseguir el equilibrio entre intentar conseguir todos los estudios relevantes y no recuperar demasiados falsos positivos que deban ser excluidos a mano.

#### Sobre este tema:

- La definición de la estrategia de búsqueda es un tema no menor, sobretodo si se quiere evaluar de alguna forma qué tan buena es. Zhang et al. proponen una técnica sistemática de aproximación a una búsqueda "casi perfecta" utilizando una sub-búsqueda manual e iteraciones [10].

El sesgo en la publicación se refiere a que es más probable que sean publicados estudios con resultados (en algún sentido) positivos a estudios que presentan resultados negativos. Para evitar esto se puede consultar la literatura gris, como por ejemplo reportes técnicos, publicaciones rechazadas o trabajos en progreso.

#### Recomendaciones [8]

- Existen diferentes estrategias de búsqueda asociadas a diferentes tipos de criterios de búsqueda. Se debe elegir y justificar la estrategia de búsqueda más apropiada para las preguntas de investigación a responder.
- Se deben utilizar diferentes fuentes electrónicas; seguramente no exista una única fuente que pueda proporcionar todos los estudios primarios.
- Actualmente, los motores de búsqueda de ingeniería de software no están diseñados para soportar las revisiones sistemáticas de literatura. Así es que muchas veces las búsquedas quedan diseñadas dependientes de la fuente.
- Las búsquedas deben aplicarse sobre los mismos campos (esto es: título, *abstract*, etc) y utilizando los mismos filtros para todos los artículos en todas las fuentes. De todas maneras, la cadena de búsqueda debe adaptarse para poder ejecutarse en las distintas bibliotecas digitales.

### 3.4.2. Seleccionar los estudios primarios

Una vez que se han obtenido los estudios primarios potenciales es necesario evaluar cabalmente su relevancia. Para esto se utilizan criterios de inclusión y exclusión, los primeros definen que estudios se deben incluir como relevantes mientras que los últimos se aplican sobre los estudios seleccionados o sobre la lista inicial para remover estudios irrelevantes.

Los criterios de selección de estudios pretenden identificar los estudios primarios que proporcionan evidencia directa acerca de las preguntas de investigación. A fin de reducir la probabilidad de sesgo, los criterios de selección deben ser decididos durante la definición de protocolo, aunque tal vez puedan ser refinados durante el proceso de búsqueda.

Los criterios de inclusión y exclusión deberían estar basados en las preguntas de investigación; y probados para asegurar que son interpretados de forma confiable y que clasifican los estudios correctamente.

#### Recomendaciones [8]

- En general, los resúmenes estándar en IT e ingeniería de software son demasiado pobres para poder seleccionar estudios primarios. Así que es recomendable revisar también las conclusiones.

### 3.4.3. Evaluar la calidad de los estudios

El objetivo de esta etapa es analizar y evaluar la calidad de cada estudio seleccionado a fin de determinar su inclusión o no en el proceso de extracción de datos y reporte de resultados de la revisión. En general, el propósito de la evaluación de la calidad es asegurar que los hallazgos de un estudio son relevantes y no sesgados.

Ante la dificultad inicial de contar con un acuerdo en la definición de "calidad" de un estudio Kitchenham [5] sugiere que la calidad se refiere a la medida en que el estudio minimiza el sesgo y maximiza tanto la validez interna y como la externa. A continuación se detallan estos conceptos:

- Sesgo (o error sistemático): Tendencia a producir resultados que se apartan sistemáticamente de los resultados "verdaderos".
- Validez interna: El grado en que el diseño y la realización de un estudio tienden a prevenir el error sistemático.
- Validez externa: El grado en que los efectos observados en el estudio son aplicables fuera del estudio en sí.

En general, se llama "instrumentos de calidad" a las evaluaciones detalladas de calidad; las cuales son, generalmente, listas de verificación de los posibles factores de sesgo que es necesario evaluar en cada estudio. La mayoría de las listas de verificación de calidad incluyen preguntas para evaluar el grado en que los artículos consideran el sesgo y la validez. También, es posible considerar: ítems genéricos relacionados a características del diseño particular del estudio e ítems específicos relacionados al área de la revisión.

A modo de referencia Kitchenham [5] incluye una lista bastante amplia de preguntas para estudios cuantitativos y otra para estudios cualitativos. También indica la alternativa de utilizar una escala de medida para cada ítem en las ocasiones en las cuales una respuesta de Si/No puede resultar engañosa.

Es importante que los investigadores no solo definan en el protocolo el instrumento de calidad sino que también especifiquen cómo será usada la información de calidad. Las opciones son las siguientes:

- Para asistir en la selección de estudios primarios. En este caso, la información de calidad se utiliza para construir criterios de inclusión/exclusión detallados.
- Para asistir en la síntesis y análisis de los datos. En este caso la información de calidad es usada para identificar subconjuntos de los estudios primarios e investigar si las diferencias de calidad están asociadas con diferentes resultados de los estudios primarios.

Es posible contar con ambos tipos de información de calidad en una misma revisión sistemática.

### Limitaciones de la Evaluación de la Calidad [5]

A menudo el informe de los estudios primarios es pobre, por lo que quizás no sea posible determinar la forma de evaluar un criterio de calidad. Es tentador suponer que porque algo no se encuentra reportado, no se hizo. Esta suposición puede ser incorrecta. Los investigadores deberían intentar obtener más información por parte de los autores del estudio.



Es posible identificar un análisis estadístico inadecuado o inapropiado, pero sin acceso a los datos originales no es posible corregirlo. Es común que los datos sean confidenciales y que no puedan estar disponibles. En algunos casos, los involucrados pueden negarse a compartir sus datos con otros investigadores debido a que desean continuar publicando sobre el tema.

#### Recomendaciones [8]

- Existe un énfasis en la necesidad de evaluar la calidad de los estudios primarios. Sin embargo, esto depende del tipo de revisión sistemática de literatura que se lleve a cabo. Por ejemplo, es común que para algunos tipos de revisión ya es un criterio de calidad suficiente que los estudios primarios hayan sido publicados en alguna revista académica.
- Es importante asegurar que la evaluación de calidad sea utilizada en las actividades siguientes de agregación y análisis de datos.

#### 3.4.4. Extraer datos

El objetivo de esta etapa es el diseño de formularios de extracción de datos para que los investigadores puedan registrar adecuadamente la información que obtienen de los estudios primarios. A fin de reducir posibles sesgos, se recomienda definir los formularios de extracción de datos y realizar pruebas piloto durante la definición del protocolo de la revisión.

Los formularios de extracción de datos deben ser diseñados para recolectar toda la información necesaria para contestar las preguntas de la revisión así como los criterios de calidad de los estudios. Si los criterios de calidad van a ser utilizados para identificar criterios de inclusión/exclusión, entonces se requieren formularios separados (ya que la información debe ser recolectada previo a la extracción de datos).

En general un formulario de extracción de datos debería incluir las siguientes partes (secciones) [3]: información de la extracción (responsable y fecha de la extracción, responsable de verificar), información general de los estudios (identificador del estudio, título y detalles de publicación), preguntas para responder las preguntas de la revisión, preguntas para evaluar la calidad de los estudios y resumen de los datos.

Es importante no incluir en la síntesis de una revisión sistemática múltiples publicaciones de los mismos datos ya que informes duplicados podrían afectar seriamente el sesgo de cualquier resultado. Cuando hay publicaciones duplicadas conviene usar la más completa.

#### Recomendaciones [8]

- Cuando sea necesario revisar un gran número de artículos puede resultar útil que un lector extraiga los datos y otro los valide.
- Si es necesario manipular datos se debe describir y justificar las manipulaciones realizadas.

- En caso de que los datos de un estudio no estén completos se puede escribir a los autores para conseguirlos.

### 3.4.5. Sintetizar los datos

La síntesis de los datos involucra recolectar y resumir los resultados incluidos en los estudios primarios. En general, hay dos tipos de síntesis de datos: síntesis descriptiva (o narrativa) y síntesis cuantitativa [3].

Con el fin de lograr conclusiones confiables la síntesis debería considerar la fuerza de la evidencia, explorar la consistencia y discutir inconsistencias. El enfoque debería estar definido en el protocolo y se determina por el tipo de las preguntas de investigación, pero también por el tipo de estudios disponibles y por la calidad de los datos.

Sin importar el tipo de síntesis, se debería comenzar con la creación de un resumen de los estudios incluidos. En general se presenta una tabla con detalles importantes como son el tipo, intervenciones, número y característica de los participantes, resultados, etc. También se presentan (en la misma o en otra tabla) elementos de la calidad de los estudios y riesgos de sesgo. Además, este proceso descriptivo debería ser explícitamente riguroso y ayudar a concluir que los estudios son similares y confiables para su síntesis [3].

La información extraída sobre los estudios (por ejemplo: intervención, población, contexto, tamaño de la muestra, resultados, calidad del estudio) debería presentarse tabulada de una manera consistente con las preguntas de la revisión. Las tablas deberían estar estructuradas para resaltar similitudes y diferencias entre los resultados de los estudios.

Es importante identificar si los resultados de los estudios son consistentes unos con otros (o sea, homogéneos) o inconsistentes (o sea, heterogéneos). Los resultados quizás puedan ser tabulados para mostrar el impacto de potenciales fuentes de heterogeneidad (por ejemplo: tipos de estudio, calidad del estudio o tamaño de la muestra).

Kitchenham sugiere realizar un análisis de sensibilidad para determinar si los estudios de baja calidad tienen impacto significativo sobre los resultados de la síntesis. El análisis de sensibilidad también puede realizarse sobre diferentes subconjuntos de los estudios primarios para determinar la robustez de los resultados.

Sobre este tema:

- La síntesis de los datos es un tema con bastantes componentes estadísticos y matemáticos. Inicialmente se pensaba que no iba ser posible realizar síntesis cuantitativas en ingeniería de software (debido a la calidad de los datos) y solo se explicaba en detalle la síntesis narrativa [5]. En [11] plantean una revisión terciaria estudiando los tipos de síntesis realizadas hasta el momento y se percibe cierto incremento en la formalidad al realizar síntesis de datos. El estudio es muy completo y es bastante ilustrativo, sobre todo porque hace una introducción bastante profunda del tema, así como una lista de los tipos de síntesis disponibles y recomendaciones sobre su uso. Para complementar este tema, se puede consultar [12] donde se presenta una visión sobre la síntesis particular de casos de estudios. Un estudio bastante más exhaustivo

sobre síntesis de datos y meta-análisis puede leerse en [7].

### 3.5. Informar sobre la revisión

La fase final de una revisión sistemática involucra redactar los resultados de la revisión y difundir los resultados a potenciales interesados.

Sobre este tema:

- Escribir artículos científicos, o material publicable en general, es difícil al principio y es necesario completar cierta curva de aprendizaje. Como siempre la práctica hace al maestro. Un libro muy bueno para comenzar es uno publicado por Björn [13], aunque nace en el área de la medicina es simple y sus recomendaciones aplicables en general.

#### 3.5.1. Especificar los mecanismos de difusión

Es importante comunicar los resultados de la revisión sistemática. Por esto es que algunas guías recomiendan planificar la estrategia de difusión durante la etapa de puesta en marcha de la revisión (si existe) o cuando se prepara el protocolo de revisión sistemática.

La mayoría de los académicos asumen que la difusión se trata de publicar en revistas académicas o presentar en conferencias. Sin embargo, si se quiere que los resultados tengan influencia sobre la práctica profesional seguramente sean necesarios otros medios de comunicación, en particular: Revistas con orientación profesional, prensa popular o especializada, folletos breves con resúmenes, posters, páginas webs, entre otros.

#### 3.5.2. Formatear el informe principal

Kitchenham [5] sugiere la siguiente estructura y contenido para informes de revisiones sistemáticas.

Sección	Subsección	Alcance / Comentarios
Título		Debería ser corto pero informativo, así como basado en las preguntas a responder por la revisión. Para revistas académicas, en general se incluye que se trata de una revisión sistemática.
Autoría		
Resumen Ejecutivo o Estructurado	Contexto	La importancia de las preguntas de investigación que aborda la revisión.
	Objetivos	Las preguntas contestadas por la revisión sistemática.
	Métodos	Fuentes de datos, selección de estudios, evaluación de calidad y extracción de datos.
	Resultados	Principales hallazgos incluyendo meta-análisis y análisis de sensibilidad.
	Conclusiones	Implicancias a la práctica y futura investigación.
Antecedentes		Justificación de la necesidad de la revisión. Resumen de revisiones previas.  <i>Descripción de la técnica de ingeniería de software que está siendo investigada así como su importancia potencial.</i>
Preguntas de la		Se debe especificar cada una de las preguntas

revisión		de la revisión.
Métodos de la revisión	Fuentes de datos y estrategia de búsqueda	<i>Estos datos se deberían basar en el protocolo original incluyendo información de los cambios que haya sufrido.</i>
	Selección de estudios	
	Evaluación de calidad de los estudios	
	Extracción de datos	
	Síntesis de los datos	
Estudios Incluidos y Excluidos		<p>Criterios de inclusión y exclusión. Lista de los estudios excluidos, así como las razones para la exclusión.</p> <p><i>Los criterios de inclusión y exclusión de los estudios a veces son representados como un diagrama de flujo, ya que los estudios pueden ser excluidos en distintas etapas de la revisión.</i></p>
Resultados	Hallazgos	<p>Descripción de los estudios primarios. Resultados de cualquier resumen cuantitativo. Detalles de cualquier meta-análisis.</p> <p><i>Se debería incluir resúmenes no cuantitativos de los estudios de forma de presentarlos en forma tabular. Los resúmenes cuantitativos se deberían presentar en tablas y gráficos.</i></p>
	Análisis de sensibilidad	
Discusión	Hallazgos principales	<i>Esto debe corresponder con los hallazgos presentados en la sección de resultados.</i>
	Fortalezas y debilidades	<p>Fortalezas y debilidades de la evidencia incluida en la revisión. Relación a otras revisiones, considerando particularmente cualquier diferencia en calidad y resultados.</p> <p><i>Una discusión de la validez de la evidencia considerando el sesgo en la revisión sistemática permite que el lector pueda evaluar la confianza que puede depositar en la evidencia recolectada.</i></p>
	Significado de los hallazgos	<p>Dirección y magnitud de los efectos observados en los estudios resumidos. Aplicabilidad (generalización) de los hallazgos.</p> <p><i>Dejar en claro en qué medida los resultados implican causalidad discutiendo el nivel de evidencia. Discutir todos los beneficios, efectos adversos y riesgos. Discutir variaciones en los efectos y sus razones.</i></p>
Conclusiones	Recomendaciones	<p>Implicaciones prácticas para el desarrollo de software. Preguntas sin responder e implicaciones para investigación futura.</p>
Agradecimientos		Todas las personas que colaboraron en la investigación pero no califican en los criterios de autoría.
Conflictos de intereses		<i>Se debe declarar cualquier interés secundario por parte de los investigadores (por ejemplo: un interés financiero en la tecnología que está siendo evaluada).</i>
Referencias y Anexos		<i>Los anexos pueden ser utilizados para: listar estudios incluidos y excluidos, documentar detalles de la estrategia de búsqueda, y listar datos sin pulir de los estudios incluidos.</i>
Tabla 2 – Índice sugerido para informe de una revisión sistemática de literatura.		

### 3.5.3. Evaluar el informe

Los informes de la revisión sistemática suelen ser evaluados en el marco de su difusión: un artículo en una revista académica será revisado previo a su publicación, expertos revisarán una tesis de doctorado. Si el protocolo fue revisado por un grupo de expertos previo a la ejecución de la revisión entonces se recomienda que el mismo grupo revise el informe final. El proceso de evaluación puede utilizar las listas de verificación de calidad para revisión sistemáticas similares a las vistas en la sección.

## 4. Estudios de Mapeo

Los estudios de mapeo (o de alcance) son un tipo de revisión sistemática de literatura, pero a diferencia de las revisiones sistemáticas convencionales su propósito es encontrar y clasificar estudios primarios dentro de un tópico específico [14].

### 4.1. Diferencias entre un estudio de mapeo y una SLR

La diferencia principal es que una revisión sistemática convencional intenta agregar los estudios primarios en términos de los resultados de la investigación e investiga si son consistentes o contradictorios. En cambio, un estudio de mapeo aspira solamente a clasificar la literatura relevante y a clasificar los estudios con respecto a categorías definidas. A continuación, se presentan las diferencias entre los estudios de mapeo y las revisiones sistemáticas convencionales [15].

Proceso de la SLR	Estudio de Mapeo	SLR
Preguntas de investigación	Generales — relacionadas a las tendencias en investigación. Qué subtópicos se manejan, qué investigadores, cuánta actividad, qué tipos de estudios, etc.	Específicas — relacionadas a los resultados de los estudios empíricos. De la forma: ¿es la tecnología/método A mejor o no que el B?
Proceso de búsqueda	Definido por el área temática.	Definido por las preguntas de investigación.
Estrategia de búsqueda	No muy estricta si sólo son de interés las tendencias en investigación.	Extremadamente estricta — todos los estudios relevantes deben ser encontrados.
Evaluación de calidad	No es esencial.	Es importante para asegurar que los resultados se basan en la evidencia de mejor calidad.
Resultados	Conjunto de artículos relacionados a un área temática y su clasificación según varias categorías.	Respuestas a preguntas de investigación específicas, posiblemente con calificadores (por ejemplo: los resultados aplican únicamente a novatos).

Tabla 3 – Resumen de diferencias entre las revisiones sistemáticas y los estudios de mapeo.

## 4.2. Esquemas de clasificación

Uno de los principales focos al realizar estudios de mapeo es la creación de un esquema de clasificación. Aunque si bien pueden tenerse ideas iniciales, su construcción finaliza cuando se dispone de los datos de los estudios relevantes encontrados.

En un estudio de mapeo en general se trabaja con uno o más esquemas de clasificación a fin de categorizar los estudios primarios encontrados, por ejemplo, identificando a qué país pertenece o estableciendo de qué tópico trata a partir de una lista de temas. Esto puede implicar una sola dimensión (por ejemplo clasificar según país de publicación) o varias dimensiones (por ejemplo clasificar según país, pero también según tema y fuente de publicación).

Para elegir un esquema de clasificación se puede consultar investigación previa o bibliografía relevante antes de comenzar el estudio de mapeo. En general, es buena práctica validar el esquema revisando un subconjunto de las publicaciones que serán cubiertas mediante el estudio previo a realizarlo.

Si es necesario la construcción de un esquema de clasificación es posible utilizar, por ejemplo, un *keywording* de los *abstracts* [16], esto implica la extracción de los términos clave de los resúmenes de los artículos encontrados. Esto puede ser bueno para algunos casos, pero en general conviene tener una noción básica de aspectos de clasificación y construcción de vocabularios controlados (que es el concepto real asociado a los esquemas de clasificación). Para esto se pueden consultar la guía de Zeng [17].

Al comenzar a trabajar con estudios de mapeo es difícil imaginar que esquemas de clasificación pueden resultar útiles. Algunos de las siguientes ideas pueden ayudar:

- Fechas de publicación — Una buena dimensión a estudiar es la fecha de publicación, en general permite analizar la evolución de los demás dimensiones a través del tiempo. En general se extrae el año solamente, pero para estudios particulares podría utilizarse el mes también.
- País de publicación — Muchas veces es de interés clasificar los estudios primarios por país de publicación.
- Fuente — También puede ser de interés clasificar los artículos por fuente de publicación.
- Métodos de investigación — Un aspecto muy interesante al evaluar artículos es determinar que método de investigación utilizan los autores. Pueden encontrarse distintas manera de realizar esta clasificación, pero está bastante generalizado el uso del esquema planteado por Wieringa [23]. Para cada método de investigación Wieringa presenta criterios de evaluación que pueden ser utilizados, por ejemplo, para crear criterios de evaluación de la calidad. En la Tabla 4 se muestra un resumen de la clasificación de Wieringa.

<b>Categoría</b>	<b>Descripción</b>
Evaluation research	Se investiga la aplicación práctica de una técnica. Se incluye cómo se implementa la técnica y cuáles son sus consecuencias (ventajas y desventajas). En general la técnica no es novedosa pero su aplicación práctica sí.
Proposal of solution	Se propone una solución a un problema y se argumenta su relevancia sin una validación en toda regla. La solución propuesta debe ser novedosa o por lo menos una extensión significativa de una técnica ya existente.
Validation research	Este tipo de artículos investigan las propiedades de una solución propuesta pero que aún no ha sido implementada. Posibles métodos para realizar esta validación pueden ser: experimentos, simulaciones, prototipos, análisis matemático, etc.
Philosophical papers	Explican una manera nueva de ver las cosas, un nuevo marco conceptual, etc.
Opinion papers	Contienen la opinión del autor sobre si algo está mal o bien, o cómo se hacen las cosas, etc.
Personal experience papers	En este tipo de artículos se explican experiencias personales. Su foco es el qué y no el porqué. La evidencia presentada en este tipo de artículos en general es anecdótica.
Tabla 4 – Esquema de clasificación según métodos de investigación [23].	

Sobre este tema:

- En [16] brindan una explicación bastante completa del proceso de creación que utilizaron en un caso de ejemplo.
- En [14] se presenta una comparación de varios estudios de mapeo y se comentan características de este tipo de revisiones.
- Otras referencias a los vocabularios controlados son: [18] y [19].

#### 4.3. Presentación de resultados

La presentación de resultados de los estudios de mapeo es un aspecto significativo. Mientras que en las revisiones sistemáticas de literatura los resultados, en general, pueden sintetizarse de alguna forma (cualitativa o cuantitativa), en los estudios de mapeo esto puede ser que no resulte tan fácil de realizar. Por este motivo, es posible que se opte por tablas o visualizaciones gráficas de los resultados.

Como el propósito de un estudio de mapeo está fuertemente ligado a la clasificación de los estudios primarios es frecuente ver gráficos donde se presenta la distribución de las publicaciones para una dimensión (una categoría, se pueden utilizar gráficos de torta u otros) o varias dimensiones (cruzando dos o tres categorías, por ejemplo en un gráfico de burbujas). Se muestran algunos ejemplos a continuación en las Figuras 1, 2 y 3.

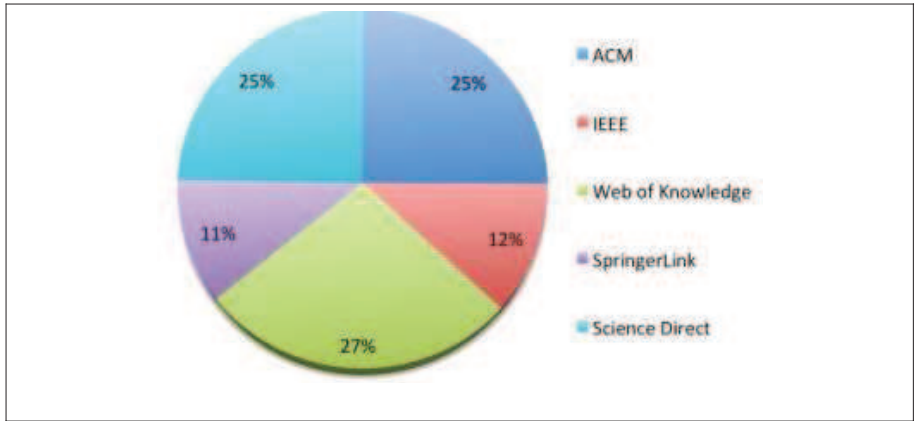


Figura 1 – Porcentaje de publicaciones por fuente bibliográfica [20] — Ejemplo de presentación de cantidad de publicaciones según una dimensión.

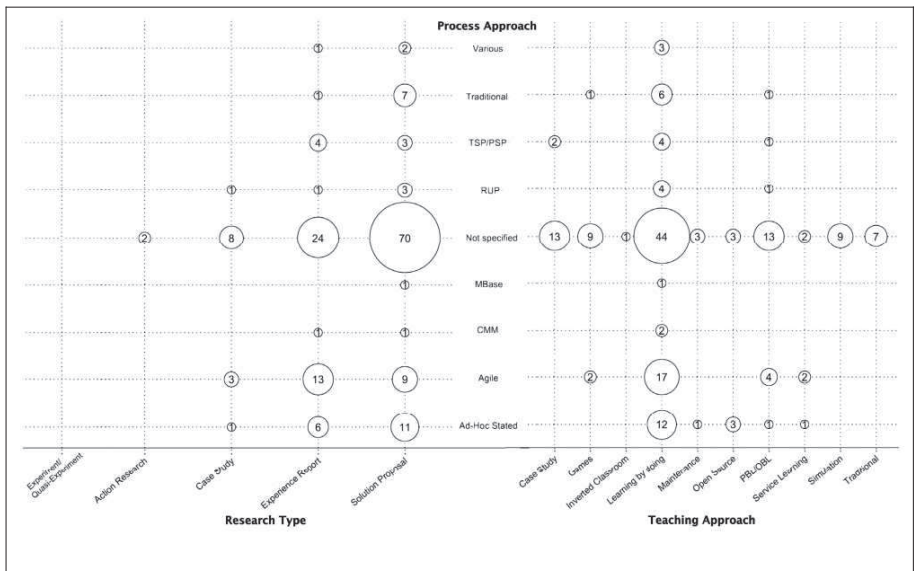


Figura 2 – Tipo de investigación versus Enfoque del proceso versus Enfoque de enseñanza [20] — Ejemplo de gráfico mostrando cantidad de publicaciones según 3 dimensiones (notar que en realidad se cruzan de a dos).



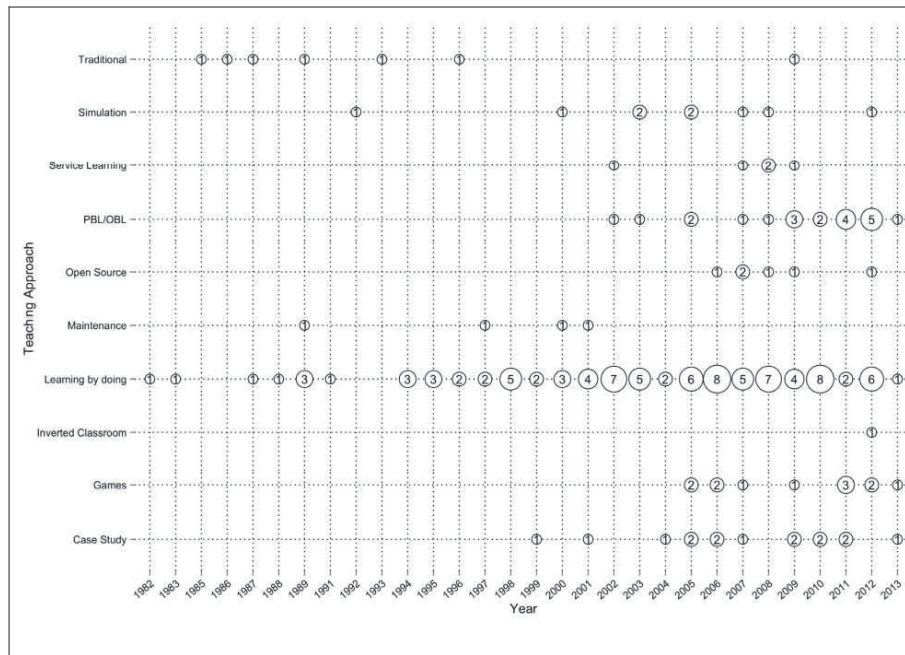


Figura 3 – Evolución de los enfoques de enseñanza a lo largo del tiempo [20] — Ejemplo de presentar la evolución de una dimensión a lo largo del tiempo.

## 5. Revisiones Terciarias

En un dominio en el cual ya exista un número de revisiones sistemáticas puede ser posible conducir una revisión terciaria (una revisión sistemática de revisiones sistemáticas) para contestar preguntas más amplias. Una revisión terciaria utiliza exactamente la misma metodología que las revisiones sistemáticas de literatura estándares.

Sobre este tema:

- A modo de ejemplo o para entender aspectos avanzados sobre revisiones sistemáticas en ingeniería de software se pueden consultar revisiones terciarias en ese tema en [21], [22], [23] y [24].

## 6. Anotaciones sobre librerías digitales y motores de búsqueda

El objetivo de esta sección es proveer información relacionada a los motores de búsqueda y librerías digitales que son utilizados para realizar investigaciones basadas en evidencia en ingeniería de software. La intención es que se actualice de forma periódica con nueva información recabada.

A modo de caso de estudio se presenta una búsqueda como ejemplo. El propósito de este caso es obtener *Evidencia empírica sobre la inclusión de los métodos formales en los currículos de informática y computación*. La cadena de búsqueda base se muestra en la Tabla 5. Se incluyen ejemplos de como utilizar las distintas bibliotecas para realizar la búsqueda adaptando cuando sea necesario nuestra cadena de búsqueda.

("formal method") AND ("curriculum" OR "curricula" OR "syllabus" OR "education" OR "course of study" OR "career" OR "degree" OR "training") AND ("software engineering" OR "computer science" OR "Computer Engineering" OR "informatics" OR "information technology" OR "information system")

Tabla 5 – Cadena de búsqueda de ejemplo. *Query string* para “Evidencia empírica sobre la inclusión de los métodos formales en los currículos de informática y computación”.

## 6.1. ACM Digital



ACM DIGITAL LIBRARY

SIGN IN SIGN UP

Full text of every article ever published by ACM and bibliographic citations from major publishers in computing.

- Using the [ACM Digital Library](#)
- For Consortia Administrators

Announcements

**Digital Library Training Sessions**

Join us for our [DL Weekly Online Training Sessions](#)

**ACM BOOKS** a dynamic new series of advanced level books in computer science, published by ACM in collaboration with Morgan & Claypool Publishers.

[learn more about the program](#) | [check out the books](#)

The ACM DL app for Android, iOS and Windows is now available (and free).

As an ACM DL user:

1. [Sign in or Sign up](#) to the ACM DL with your ACM Web Account now.
2. Install the ACM DL app onto your device, (see below).
3. Sign in to the ACM DL app with your ACM Web Account.

**DONE**—You now have access to the ACM DL on your mobile device.

GET IT ON Google play

Download on the App Store

Download from Windows Phone Store

**Advanced Search**

**Browse the ACM Publications:**

- Journals/Transactions
- Magazines
- Proceedings
- ACM Books

**Browse the Special Interest Groups:**

- Special Interest Groups (SIGs)

**Browse the Conferences:**

- Recent and Upcoming Conferences
- Conference Listing

**Browse the Special Collections:**

- eBooks available to ACM Members
- ACM International Conference Proceeding Series (ICPS)
- Classic Book Series
- ACM Oral History interviews
- ACM Curricula Recommendations
- NSF Workshop Reports

**Browse the Publications by Affiliated Organizations**

**Browse all literature by type** [select a type] ▾

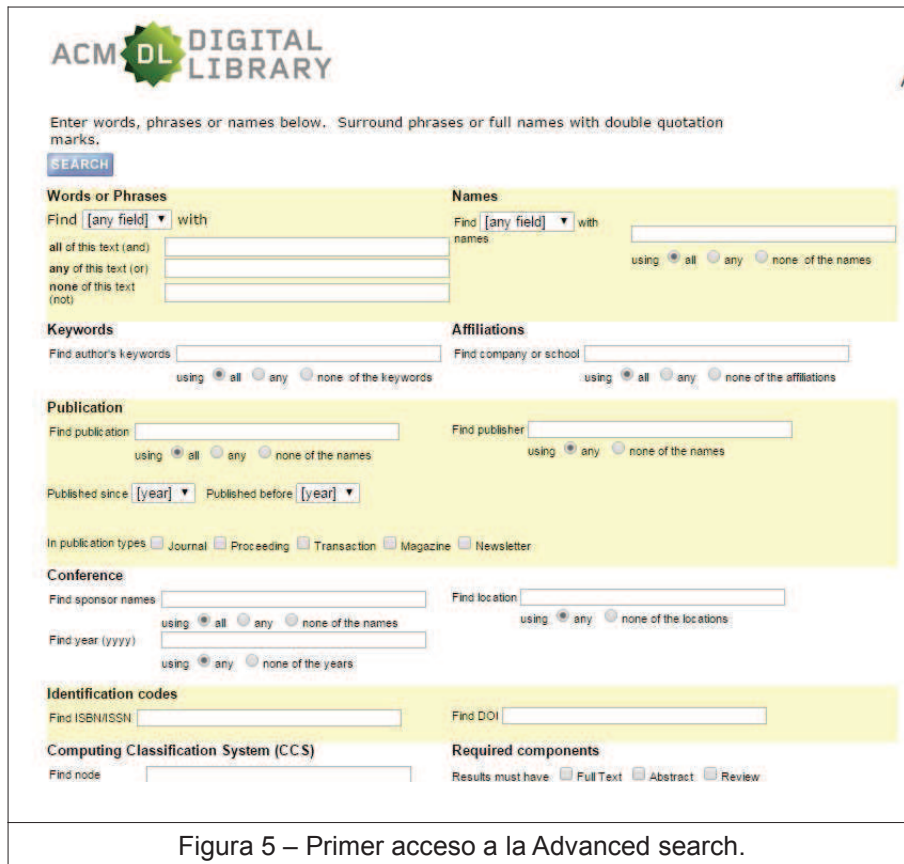
**Browse all literature by Publisher**

**Browse by the ACM Computing Classification System**

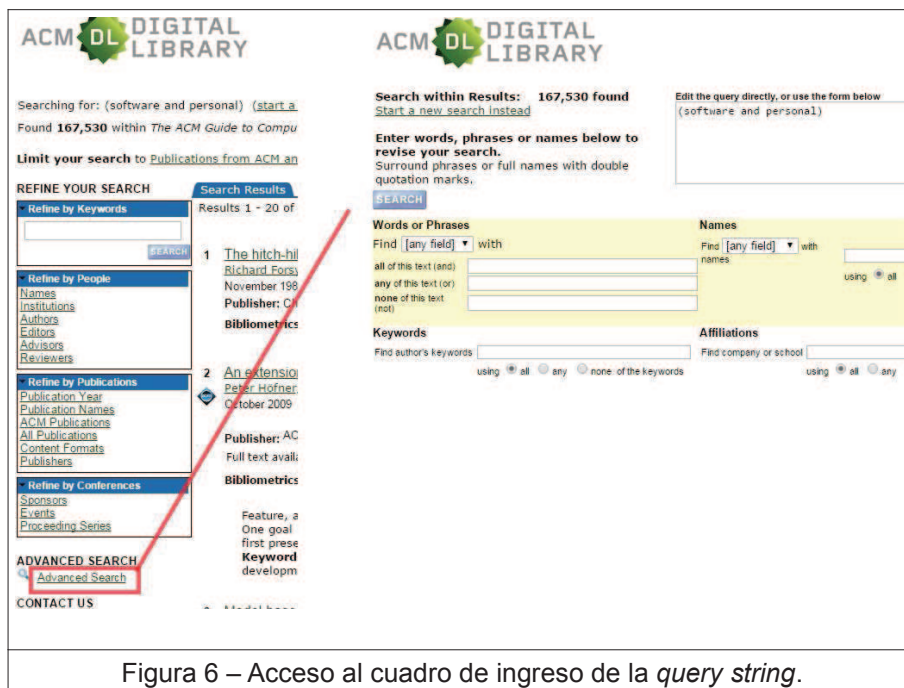
[The Encyclopedia of Computer Science](#)  
"...is the definitive reference in computer science and technology..."

Figura 4 – Página de acceso a ACM Digital.

ACM Digital indexa *journals* y *proceedings* de conferencias publicados por la ACM así como por organizaciones afiliadas. La biblioteca tiene una sección de Advanced search que se accede mediante el link en la esquina superior derecha (Figura 4). En principio es poco intuitivo el uso ya que no presenta un campo para ingresar una *query string* ni una guía de la sintaxis (Figura 5).



Se pueden utilizar los campos del formulario de búsqueda y así generar la query que se necesite. Acepta *wildcards*, pero los operadores + y - son reservados. Es posible disponer del cuadro de ingreso de la *query string*, pero para esto se debe realizar una consulta cualquiera y luego acceder al link *Advanced search* desde los resultados de la búsqueda. Esto se muestra en la Figura 6.



En la Tabla 6 se incluye la cadena de búsqueda del caso de estudio adaptada para esta librería.

```
(Title:"formal method" or Abstract:"formal method") AND ((Title:curriculum or Abstract:curriculum) OR (Title:curricula or Abstract:curricula) OR (Title:syllabus or Abstract:syllabus) OR (Title:education or Abstract:education) OR (Title:"course of study" or Abstract:"course of study") OR (Title:career or Abstract:career) OR (Title:degree or Abstract:degree) OR (Title:training or Abstract:training)) AND ((Title:"software engineering" or Abstract:"software engineering") OR (Title:"computer science" or Abstract:"computer science") OR (Title:"Computer Engineering" or Abstract:"Computer Engineering") OR (Title:"informatics" or Abstract:"informatics") OR (Title:"information technology" or Abstract:"information technology") OR (Title:"information technology" or Abstract:"information technology") OR (Title:"information system" or Abstract:"information system"))
```

Tabla 6 – Cadena de búsqueda adaptada para ACM Digital.

**Atención:** Si bien Scopus indexa ACM, utilizando ambos buscadores se encontró que algunos artículos que aparecen en ACM no eran devueltos por Scopus.

## 6.2. IEEEXplore

Esta biblioteca indexa todas las publicaciones de la IEEE. Tiene una sección de *Advanced search* (Figura 7) bastante intuitiva y además cuenta con la opción de *Command search* desde la cual se puede ingresar directamente una *query string* (Figura 8).

Es importante notar que los resultados se muestran por defecto ranqueados por “relevancia”, donde la relevancia se calcula teniendo en cuenta frecuencia y proximidad de los términos de búsqueda en el resumen y en el registro de citas, así como el largo del resumen y del registro de citas.



Figura 7 – *Advanced search* de la biblioteca IEEE Xplore.

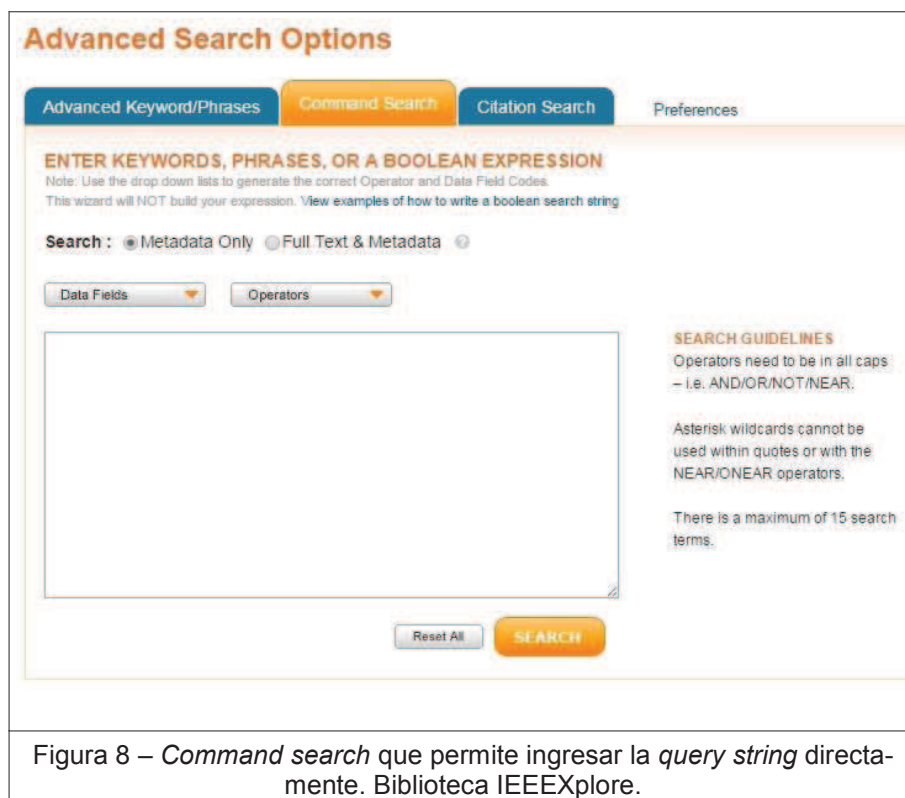


Figura 8 – *Command search* que permite ingresar la *query string* directamente. Biblioteca IEEE Xplore.

Se puede buscar por metadata, la cual incluye el *abstract*, *indexing terms* y *bibliographic citation information* (*article title*, *author*, *publication title*, etc.). Esta opción es bastante útil por su simplicidad. En la Ta-



bla 7 se muestra la cadena de búsqueda del caso de estudio adaptada utilizando esta opción.

```
("formal method") AND ("curriculum" OR "curricula" OR "syllabus" OR "education" OR "course of study" OR "career" OR "degree" OR "training") AND ("software engineering" OR "computer science" OR "Computer Engineering" OR "informatics" OR "information technology" OR "information system"))
```

Tabla 7 – Cadena de búsqueda adaptada para IEEEExplore, buscando sobre la metadata.

Para buscar por *abstract* basta con poner el prefijo "Abstract": a cada término (por ejemplo: "Abstract":*formal method*). Lo mismo para *title* con el prefijo "Document Title".

Sí la sintaxis de la búsqueda es incorrecta, por ejemplo por omitir un operador, no indica un mensaje de error, sino que muestra que no se encontró ningún resultado. De todas formas hay una guía de búsqueda que advierte que se tienen que utilizar operadores y tiene varios ejemplos. Dicha guía también advierte que se pueden utilizar como máximo 15 términos, pero no dice que tiene un límite de 40 palabras.

Si se ingresa una palabra o frase pero sin comillas, IEEE busca también variantes de esa frase o palabra. Por ejemplo palabras en plural, conjugaciones de verbos. Por ejemplo si la cadena de búsqueda es: "run", también busca, por ejemplo, por "runs", "running", etc. Si la palabra está en inglés británico, también la busca en inglés americano. Por ejemplo si se ingresa la palabra "color", IEEE también busca por "colour".

Si se quiere que una frase o palabra aparezca exactamente en el artículo, se debe poner entre comillas. El carácter " no es reconocido por el buscador (sería lo mismo que no ingresar ninguna comilla), en cambio " si. Por ejemplo, si la cadena de búsqueda es: "formal method" AND "information technology" va a filtrar por artículos con la frase "formal method" y con la palabra "information" o "technology" (o alguna variante), devolviendo más resultados del esperado.

Hay ciertas palabras que son ignoradas por el buscador por ser generales o demasiado frecuentes. Por ejemplo, si la cadena de búsqueda es: "the computer", IEEE Explorer va a buscar por la palabra "computer". Estas palabras son: a, as, have, of, what, about, at, how, on, when, above, can, I, or, where, an, do, in, over, why, and, find, is, show, with, any, for, me, the, you, are, from, not, under, your. Si se quiere que en la búsqueda aparezcan estas palabras, la frase se tiene que poner entre comillas.

El buscador ignora algunos caracteres de puntuación. Por ejemplo si la cadena de búsqueda es "solid-state", IEEE va a buscar por "solid-state", "solid state" y "solid\_state", así como cualquier otra ocurrencia de la frase "solid state" con caracteres de puntuación entre las palabras. Si la cadena de búsqueda es: "formal method" va a devolver el resultado del artículo que contiene la frase, por ejemplo, "non-formal method".

Los caracteres especiales, incluyendo caracteres no alfanuméricos, como % , - , @, etc, son ignorados por el buscador. Los únicos caracteres especiales reconocidos por IEEE Xplore son &, +, y /. El carácter \* también es especial en el buscador.

En la sección *Advanced Search Options* hay una opción *Preferences* (también se puede acceder desde *My settings*-->*Preferences*). En la misma se pueden setear las distintas opciones de búsqueda, como por ejemplo la cantidad de resultados por página. La Figura 9 muestra las opciones de preferencias.

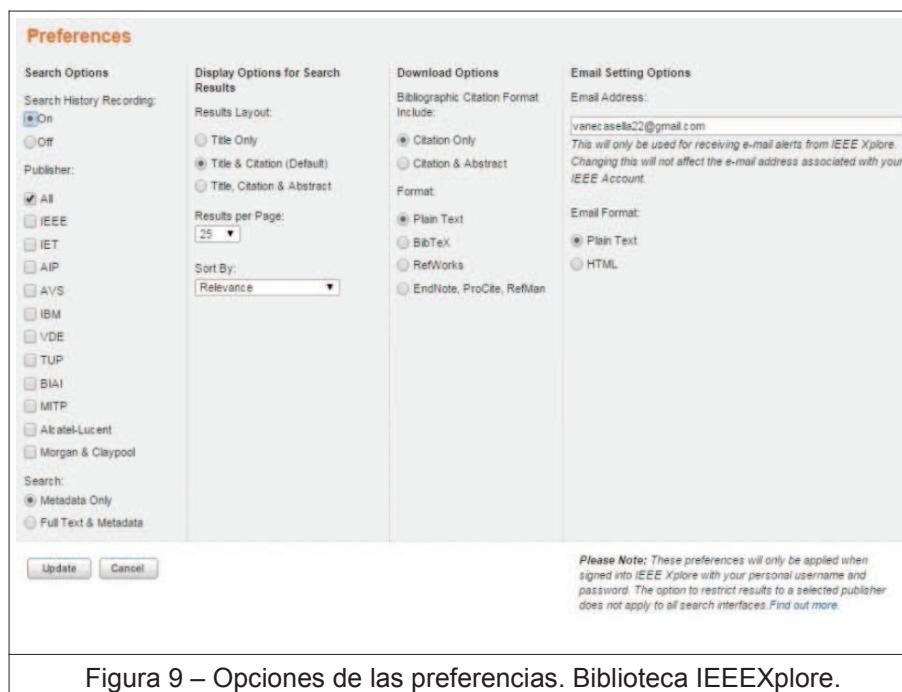


Figura 9 – Opciones de las preferencias. Biblioteca IEEEExplore.

Algo que resulta muy útil es la opción *Search history recording*, que si está en *on* guarda las búsquedas que se realizan (por defecto esta en *off*). Todas las búsquedas realizadas se pueden ver en *My Settings*-->*Search History*.

**Atención:** Si bien Scopus indexa IEEE, utilizando ambos buscadores se encontró que algunos artículos que aparecen en IEEE no eran devueltos por Scopus.

### 6.3. ScienceDirect

ScienceDirect es una biblioteca que indexa *journals* publicados por Elsevier y otros publicados por CrossRef (<http://www.crossref.org/>). Esta biblioteca tiene dos opciones avanzadas: *Advanced search* y *Expert search*. La primera (Figura 10) permiten construir de forma asistida una consulta, seleccionando una combinación de AND, OR o AND NOT. Además se pueden aplicar algunos filtros, como años, disciplinas, etc. Permite buscar por *abstract*, *title*, *keywords*, *references*, etc.

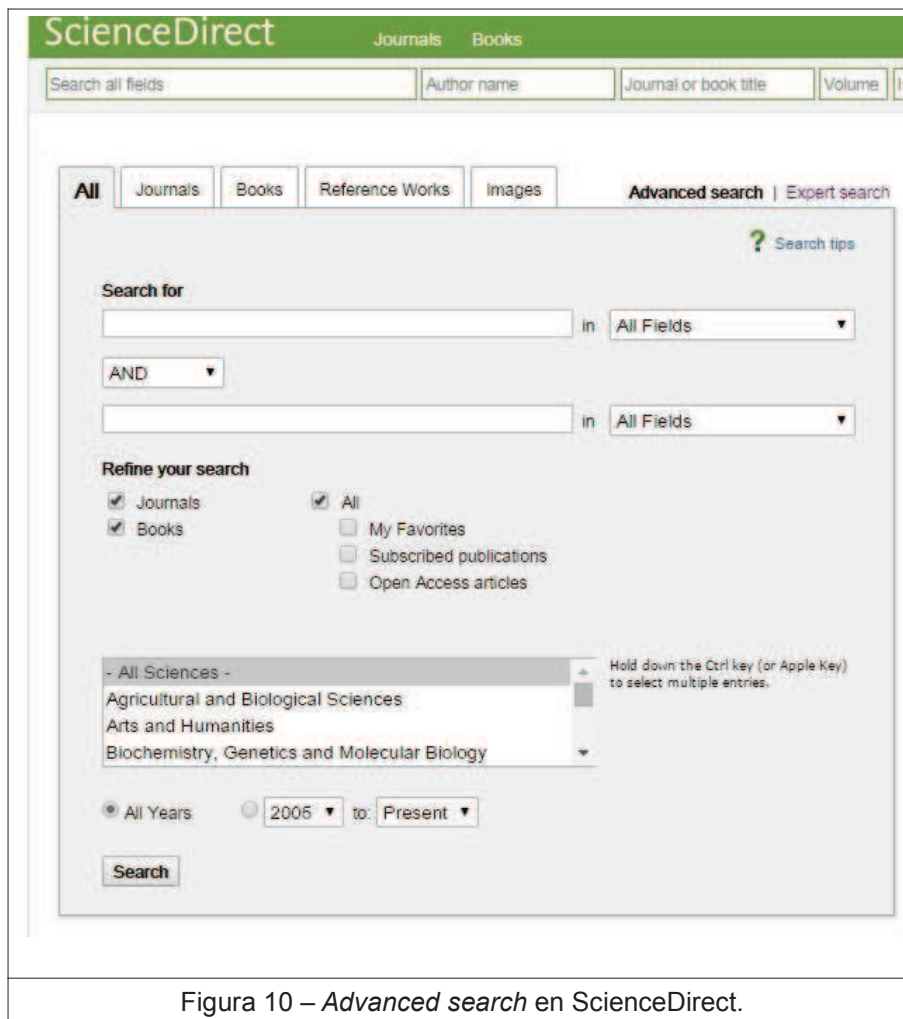


Figura 10 – *Advanced search* en ScienceDirect.

La funcionalidad de *Expert search* permite realizar una búsqueda a través de una *query*. También se pueden aplicar algunos filtros, como años, disciplinas, etc.

No hay directamente una opción para buscar por *abstract*, *title*, *keywords*, *references*, etc. Si se quiere buscar específicamente por algún campo, hay que adaptar la cadena de búsqueda. Una opción es generar una búsqueda utilizando la opción *Advanced search* (la cual te muestra la cadena de búsqueda correspondiente) para luego poder escribir la cadena de búsqueda en *Expert search* de forma adecuada. En la Tabla 8 se muestra la cadena de búsqueda del caso de estudio adaptada a esta plataforma.

```
TITLE-ABSTR-KEY("formal method") AND (TITLE-ABSTR-KEY("curriculum") OR TITLE-ABSTR-KEY("curricula") OR TITLE-ABSTR-KEY("syllabus") OR TITLE-ABSTR-KEY("education") OR TITLE-ABSTR-KEY("course of study") OR TITLE-ABSTR-KEY("career") OR TITLE-ABSTR-KEY("degree") OR TITLE-ABSTR-KEY("training")) AND (TITLE-ABSTR-KEY("software engineering") OR TITLE-ABSTR-KEY("computer science") OR TITLE-ABSTR-KEY("Computer Engineering") OR TITLE-ABSTR-KEY("informatics") OR TITLE-ABSTR-KEY("information technology") OR TITLE-ABSTR-KEY("information system"))
```

Tabla 8 – Cadena de búsqueda adaptada para ScienceDirect



Existen una opción *Search history* que por defecto está en *Turn on* y una opción *Search tips*.

**Atención:** Si bien Scopus indexa ScienceDirect, utilizando ambos buscadores se encontró que algunos artículos que aparecen en ScienceDirect no eran devueltos por Scopus.

## 6.4. Scopus

Es un motor de búsqueda que indexa diferentes *journals* y bibliotecas, la lista completa puede accederse desde <http://www.elsevier.com/solutions/scopus/content>. Tiene una sección para *Advanced search* (Figura 11).

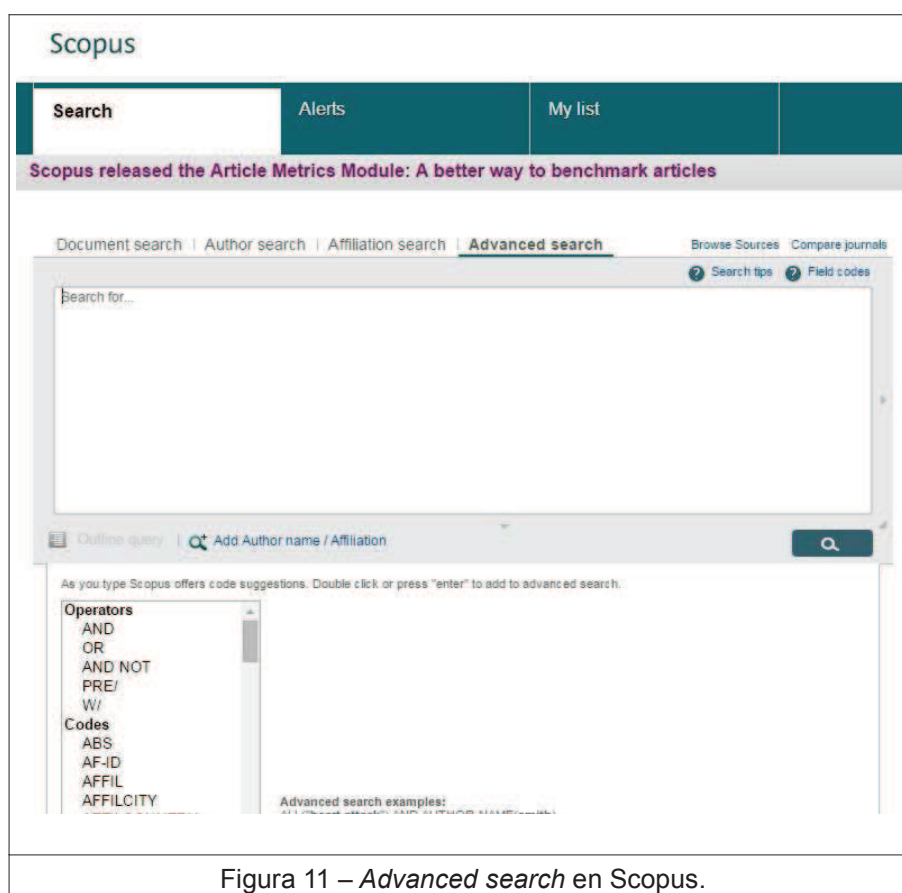


Figura 11 – *Advanced search* en Scopus.

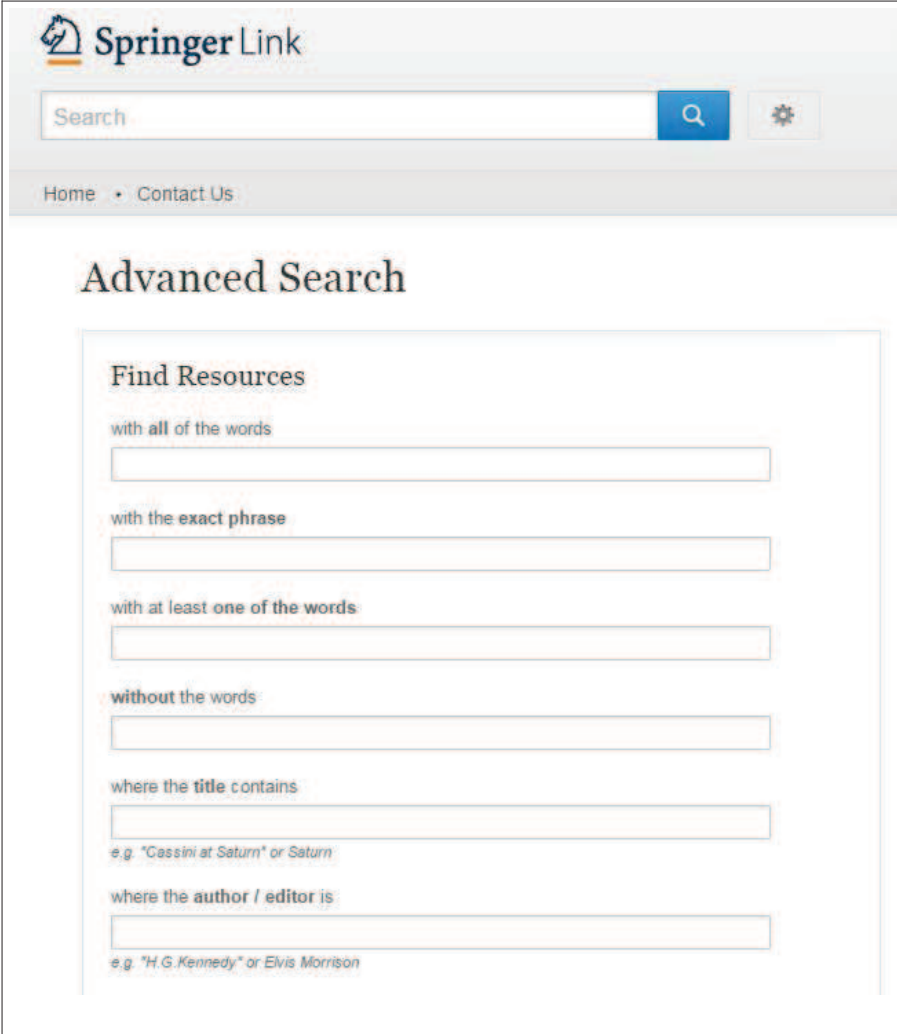
Es posible utilizar diferentes filtros para obtener resultados más acertados, por ejemplo: períodos de años, áreas temáticas, etc.

```
("formal method") AND ("curriculum" OR "curricula" OR "syllabus" OR "education" OR "course of study" OR "career" OR "degree" OR "training") AND ("software engineering" OR "computer science" OR "Computer Engineering" OR "informatics" OR "information technology" OR "information system")
```

Tabla 9 – Cadena de búsqueda adaptada para Scopus

## 6.5. SpringerLink

Esta biblioteca digital tiene una sección para *Advanced search* (Figura 11). La biblioteca busca en texto completo por lo que la cadena de búsqueda original retorna demasiados resultados. Se puede refinar específicamente la cadena para este buscador, evitando excesivas combinaciones de palabras, pero a veces es inviable. Puede ser que alcance con buscar en Scopus dependiendo de las publicaciones en las que se esté buscando.



The image shows the SpringerLink website interface for an advanced search. At the top left is the SpringerLink logo. Below it is a search bar with the placeholder text "Search", a blue search button with a magnifying glass icon, and a settings gear icon. A navigation bar below the search bar contains "Home" and "Contact Us". The main heading is "Advanced Search". Below this is a section titled "Find Resources" containing six search criteria, each with a text input field:

- with **all** of the words
- with the **exact phrase**
- with at least **one** of the words
- without** the words
- where the **title** contains  
e.g. "Cassini at Saturn" or Saturn
- where the **author / editor** is  
e.g. "H.G. Kennedy" or Elvis Morrison

At the bottom of the page, there is a caption: "Figura 12 – Advanced search en SpringerLink."

## 7. Referencias

- [1] B. A. Kitchenham, T. Dyba, y M. Jorgensen, «Evidence-based software engineering», en *Software Engineering, 2004. ICSE 2004. Proceedings. 26th International Conference on*, 2004, pp. 273-281.
- [2] T. Dyba y T. Dingsøyr, «Strength of Evidence in Systematic Reviews in Software Engineering», en *Proceedings of the Second ACM-IEEE*

*International Symposium on Empirical Software Engineering and Measurement*, New York, NY, USA, 2008, pp. 178–187.

- [3] Z. Stapić, E. G. López, A. G. Cabot, L. de Marcos Ortega, y V. Strahonja, «Performing systematic literature review in software engineering», *Proc. 23rd Cent. Eur. Conf. Inf. Intell. Syst. Varažd.*, pp. 441–447, 2012.
- [4] C. Wohlin, P. Runeson, M. Hst, M. C. Ohlsson, B. Regnell, y A. Wessln, *Experimentation in Software Engineering*. Springer Publishing Company, Incorporated, 2012.
- [5] B. Kitchenham y S. Charters, *Guidelines for performing Systematic Literature Reviews in Software Engineering*. 2007.
- [6] S. Easterbrook, J. Singer, M.-A. Storey, y D. Damian, «Selecting Empirical Methods for Software Engineering Research», en *Guide to Advanced Empirical Software Engineering*, F. Shull, J. Singer, y D. K. Sjøberg, Eds. Springer London, 2008, pp. 285-311.
- [7] J. Akers, R. Aguiar-Ibáñez, A. Baba-Akbari Sari, S. Beynon, A. Booth, J. Burch, D. Chambers, D. Craig, J. Dalton, S. Duffy, A. Eastwood, D. Fayter, T. Fonseca, D. Fox, J. Glanville, S. Golder, S. Hempel, K. Light, C. McDaid, G. Norman, B. Phillips, C. Pierce, S. Rice, A. Rithalia, M. Rodgers, F. Sharp, A. Sowden, L. Stewart, C. Stock, R. Trowman, R. Wade, M. Westwood, P. Wilson, N. Woolacott, G. Worthy, y K. Wright, *Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care*, vol. 3rd. York, UK: Centre for Reviews and Dissemination, University of York, 2009.
- [8] P. Brereton, B. A. Kitchenham, D. Budgen, M. Turner, y M. Khalil, «Lessons from applying the systematic literature review process within the software engineering domain», *J. Syst. Softw.*, vol. 80, n.º 4, pp. 571-583, abr. 2007.
- [9] J. Biolchini, P. G. Mian, y A. C. C. Natali, «Systematic Review in Software Engineering», COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, RT-ES 679/05, may 2005.
- [10] H. Zhang, M. A. Babar, y P. Tell, «Identifying relevant studies in software engineering», *Inf. Softw. Technol.*, vol. 53, n.º 6, pp. 625 - 637, 2011.
- [11] D. S. Cruzes y T. Dybå, «Research synthesis in software engineering: A tertiary study», *Inf. Softw. Technol.*, vol. 53, n.º 5, pp. 440 - 455, 2011.
- [12] D. S. Cruzes, T. Dyba, P. Runeson, y M. Host, «Case Studies Synthesis: Brief Experience and Challenges for the Future», en *Empirical Software Engineering and Measurement (ESEM), 2011 International Symposium on*, 2011, pp. 343-346.
- [13] B. Gustavii, *How to Write and Illustrate a Scientific Paper*, Second. Cambridge University Press, 2008.
- [14] D. Budgen, M. Turner, P. Brereton, y B. Kitchenham, «Using Mapping Studies in Software Engineering», en *PPIG 2008: In 20th Annual Meeting of the Psychology of Programming Interest Group*, 2008, pp. 195–204.
- [15] B. A. Kitchenham, D. Budgen, y O. P. Brereton, «The Value of Mapping Studies: A Participant-observer Case Study», en *Proceedings of the 14th International Conference on Evaluation and Assessment in Software Engineering*, Swinton, UK, UK, 2010, pp. 25–33.
- [16] K. Petersen, R. Feldt, S. Mujtaba, y M. Mattsson, «Systematic Mapping Studies in Software Engineering», en *Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering*, Swinton, UK, UK, 2008, pp. 68–77.
- [17] M. L. Zeng, *Constructing Controlled Vocabularies: A primer*. 2005.

- [18] M. R. Marques, A. Quispe, y S. F. Ochoa, «A systematic mapping study on practical approaches to teaching software engineering», en *Frontiers in Education Conference (FIE), 2014 IEEE*, 2014, pp. 1-8.
- [19] B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey, y S. Linkman, «Systematic Literature Reviews in Software Engineering - A Systematic Literature Review», *Inf Softw Technol*, vol. 51, n.º 1, pp. 7–15, ene. 2009.
- [20] B. Kitchenham, R. Pretorius, D. Budgen, O. P. Brereton, M. Turner, M. Niazi, y S. Linkman, «Systematic literature reviews in software engineering – A tertiary study», *Inf. Softw. Technol.*, vol. 52, n.º 8, pp. 792 - 805, 2010.
- [21] F. Q. B. da Silva, A. L. M. Santos, S. Soares, A. C. C. França, C. V. F. Monteiro, y F. F. Maciel, «Six years of systematic literature reviews in software engineering: An updated tertiary study», *Inf. Softw. Technol.*, vol. 53, n.º 9, pp. 899 - 913, 2011.
- [22] H. Zhang y M. A. Babar, «Systematic reviews in software engineering: An empirical investigation», *Inf. Softw. Technol.*, vol. 55, n.º 7, pp. 1341 - 1354, 2013.
- [23] R. Wieringa, N. Maiden, N. Mead, y C. Rolland, «Requirements Engineering Paper Classification and Evaluation Criteria: A Proposal and a Discussion», *Requir. Eng.*, vol. 11, n.º 1, pp. 102–107, dic. 2005.

## 8. Material complementario

En esta sección se incluyen 3 anexos seleccionados del sitio de ingeniería de software basado en evidencias – *Evidence-Based Software Engineering* (<http://community.dur.ac.uk/ebse/>).

Estos corresponden a:

- Una plantilla para revisiones sistemáticas de literatura — <http://community.dur.ac.uk/ebse/resources/templates/SLRTemplate.pdf>
- Una plantilla para estudios de mapeo — <http://community.dur.ac.uk/ebse/resources/templates/MappingStudyTemplate.pdf>
- Un ejemplo de una revisión sistemática de literatura — <http://community.dur.ac.uk/ebse/resources/studies/protocol/Agile-Methods-literature-review-protocol.pdf>

Los tres documentos no han sido modificados y se reproducen bajo la licencia de Creative Commons (<http://creativecommons.org/licenses/by-sa/2.0/uk/>).

# Template for a Systematic Literature Review Protocol

## 1. Change Record

This should be a list or table summarizing the main updates and changes embodied in each version of the protocol and (where appropriate), the reasons for these.

## 2. Background

- a) explain why there is a need for a study on this topic
- b) specify the main research question being addressed by this study
- c) specify any additional research questions that will be addressed
- d) if extending previous research on the topic, explain why a new study is needed

## 3. Search Strategy

- a) specify and justify basic strategy: manual search, automated search, or mixed
- b) for automated searches, specify search terms and compounds of these (and record results of any prototyping of the search strings)
- c) for automated searches, identify resources to be used (digital libraries and search engines)
- d) for manual searches, identify the journals and conferences to be searched
- e) specify the time period to be covered by the review and any reasons for your choice
- f) identify any ancillary search procedures, e.g. asking leading researchers or research groups, or accessing their web sites; or checking reference lists of primary studies
- g) specify how the search process is to be evaluated (e.g. against a known subset of papers; or against the results from a previous systematic review)

## 4. Selection Criteria

- a) identify the *inclusion* criteria for primary studies
- b) identify the *exclusion* criteria
- c) define how selection will be undertaken (roles of analysts)
- d) define how agreement among analysts will be evaluated
- e) define how any differences between analysts will be resolved

## 5. Study Quality Assessment

- a) specify the quality checklists to be used
- b) specify how the checklist will be evaluated
- c) define how agreement among data extractors will be evaluated
- d) define how any differences between data extractors will be resolved
- e) identify the procedures to use for applying the checklists (e.g. details inclusion/exclusion; partitioning the primary studies during aggregation or meta-analysis; explaining the results of primary studies)

## 6. Data Extraction

- a) design data extraction form (and check via a dry run)
- b) specify the strategy for extracting the data and the form (paper, on-line etc.)
- c) identify how the data extraction process is to be undertaken and validated, particularly any data that require numerical calculations, or are subjective

**7. Synthesis**

- a) specify the form of analysis to be used (e.g. narrative, tabulation, meta-analysis)
- b) assess the threats to validity (construct, internal, external), particularly constraints on the search process and deviations from standard practice

**8. Study Limitations**

Specify residual validity issues including potential conflicts of interest (i.e. that are inherent in the context of the study, rather than arising from the plan).

**9. Reporting**

Identify target audience, relationship to other studies, planned publications, authors of the publications.

**10. Schedule**

Provide time estimates for all of the major steps.

# Template for a Mapping Study Protocol

## 1. Change Record

This should be a list or table summarizing the main updates and changes embodied in each version of the protocol and (where appropriate), the reasons for these.

## 2. Background

- a) explain why there is a need for a study on this topic
- b) identify the topic that is to be 'scoped' in the study
- c) specify any research questions that will be addressed
- d) if extending previous research on the topic, explain why a new study is needed

## 3. Search Strategy

- a) specify and justify basic strategy: manual search, automated search, or mixed
- b) for automated searches, specify search terms and compounds of these (and record results of any prototyping of the search strings)
- c) for automated searches, identify resources to be used (digital libraries and search engines)
- d) for manual searches, identify the journals and conferences to be searched
- e) specify the time period to be covered by the review and any reasons for your choice
- f) identify any ancillary search procedures, e.g. asking leading researchers or research groups, or accessing their web sites; or checking reference lists of primary studies
- g) specify how the search process is to be evaluated (e.g. against a known subset of papers; or against the results from a previous systematic review or mapping study)

## 4. Selection Criteria

- a) identify the *inclusion* criteria for primary studies
- b) identify the *exclusion* criteria
- c) define how selection will be undertaken (roles of analysts)
- d) define how agreement among analysts will be evaluated
- e) define how any differences between analysts will be resolved

## 5. Data Extraction

- a) design data extraction form (and check via a dry run)
- b) specify the strategy for extracting the data and the form (paper, on-line etc.)
- c) identify how the data extraction process is to be undertaken and validated, particularly any data that require numerical calculations, or are subjective

## 6. Synthesis

- a) specify the categorization schemes to be used
- b) assess the threats to validity (construct, internal, external), particularly constraints on the search process and deviations from standard practice

## 7. Study Limitations

Specify residual validity issues including potential conflicts of interest (i.e. that are inherent in the context of the study, rather than arising from the plan).

## 8. Reporting

Identify target audience, relationship to other studies, planned publications, authors of the publications.



**9. Schedule**

Provide time estimates for all of the major steps.

# Review Protocol – Agile Software Development

Tore Dybå

## 1. Background

The concept of “Agile Software Development” has sparked a lot of interest in both industry and academia. Advocates of agile methods consider them the best thing that has happened to SE in recent years, while opponents view them as a backlash comparing them to hacking.

There is therefore a need to undertake a systematic review of the available empirical research relevant to agile software development to present a fair evaluation of the benefits and limitations by using a trustworthy, rigorous, and auditable methodology.

Thus, the purpose of this protocol is to specify the plan which the review will follow to identify, appraise and collate the evidence on the benefits and limitations of agile software development.

## 2. Objectives

The objectives of the review are to:

- Undertake a systematic review of empirical research on the benefits and limitations of agile software development.
- Select a sub-set of studies to review in-depth.
- Synthesize the evidence from these studies about the benefits and limitation of agile software development.
- Identify the research methods used.
- Identify any gaps in current research in order to suggest areas for further investigation.
- Provide a framework/background in order to appropriately position new research activities.

More specifically, we want to find out:

1. What is currently known about the benefits and limitations of agile software development?
2. What is the strength of the evidence in support of the claims regarding agile software development?
3. What are the implications of these studies for the software industry and the research community?

In addition to producing substantive findings regarding agile software development, the review also aims to advance methodology for integrating diverse study types, including ‘qualitative’ research, within systematic reviews of software engineering interventions.

## 3. Criteria for considering studies for this review

Types of studies:

Articles to be included in the review must present empirical data and pass the minimum quality threshold described below.

Types of participants:

Studies of both students and professional software developers will be included.

Types of intervention:

Inclusion of studies will not be restricted to any specific type of intervention.

Types of outcome measures:

Studies must include one of the primary outcomes of time-to-market, cost, or process or product quality to be included.

#### **4. Search strategy for identification of studies**

The search strategies will include electronic databases and hand searches of conference proceedings.

Electronic databases:

The following electronic databases will be searched: ISI Web of Science, IEEE Xplore, ACM Digital Library, Kluwer Online, ScienceDirect – Elsevier, Springer, Wiley Inter Science Journal Finder, and Compendex.

Conference proceedings:

The following conference proceedings will be hand searched for research papers: XP, XP/Agile Universe, and Agile Development Conference.

Search strategy:

The title, abstract, and keywords of the articles in the included electronic databases and conference proceedings will be searched according to the following search strategy:

1. agile AND software
2. extreme programming
3. xp AND software
4. scrum AND software
5. crystal AND software AND (clear OR orange OR red OR blue)
6. dsdm AND software
7. fdd AND software
8. feature AND driven AND development AND software
9. lean AND software AND development

#### **5. Methods of the review**

Selection of studies:

The search strategy will identify all relevant articles and these will be reviewed by two reviewers. Only studies written in the English language will be included.

Studies will be excluded if:

- (i) the study's focus, or main focus, is NOT agile software development;
- (ii) the study does NOT present empirical data.

If it is unclear from the title, abstract, and keywords whether a study conforms to these screening criteria, it will be included for a detailed quality assessment (see below).

Assessment of methodological quality:

Each study will be assessed independently by two reviewers according to 10 criteria (see Appendix 1). These criteria are informed by those proposed for the Critical Appraisal Skills Programme (CASP), in particular those for assessing the quality of ‘qualitative’ research (Greenhalgh, 2001), and by principles of good practice for conducting empirical research in software engineering (Kitchenham *et al.*, 2002).

The 10 criteria cover three main quality issues that need to be considered when appraising the studies of this research:

- *Rigour*: has a thorough and appropriate approach been applied to key research methods in the study?
- *Credibility*: are the findings well presented and meaningful?
- *Relevance*: how useful are the findings to the software industry and the research community?

Two screening criteria relate to the quality of the *reporting* of a study’s rationale, aims, and context. Each study will be assessed according to whether:

- (i) the aims and objectives are clearly reported (including a rationale for why the study was undertaken);
- (ii) there is an adequate description of the context in which the research was carried out.

A further five criteria relate to the *rigour* of the research methods employed to establish the validity of data collection tools and the analysis methods, and hence the trustworthiness of the findings. Each study will be assessed according to whether:

- (iii) the research design was appropriate to address the aims of the research
- (iv) there is an adequate description of the sample used and the methods for how the sample was identified and recruited;
- (v) any control groups were used to compare treatments
- (vi) appropriate data collection methods were used and described;
- (vii) there is adequate description of the methods used to analyse data and whether appropriate methods for ensuring the data analysis was grounded in the data.

In addition, two criteria relate to the assessment of the *credibility* of the study methods for ensuring that the findings are valid and meaningful. In relation to this, the reviewers will judge studies according to whether:

- (viii) the relationship between the researcher and participants has been adequately considered;
- (ix) the study provide clearly stated findings with credible results and justified conclusions.

The final criterion relates to the assessment of the *relevance* of the study for the software industry at large and the research community. Thus, the reviewers will judge studies according to whether:

- (x) they provide value for research or practice.

Taken together, these 10 criteria provide a measure of the extent to which we can be confident that a particular study’s findings can make a valuable contribution to this review.

Two researchers will carry out all the procedures in this section independently, and then meet to compare their assessments and establishing consensus. Any differences will be resolved by consulting a third reviewer.

Data management and extraction:

All data will be extracted by two reviewers. Differences in data extraction will be resolved by referring back to the original article and establishing consensus. A third reviewer will be consulted to help resolve differences.

When necessary the authors of the primary studies will be contacted to obtain additional information

The following data will be extracted regarding the *description* of the studies:

- (i) bibliographic reference
- (ii) type of article (e.g. journal, conference, workshop)
- (iii) research question
- (iv) study characteristics/phenomenon
- (v) study context

The following data will be extracted regarding the *findings* of the studies:

- (i) study design
- (ii) sampling
- (iii) data collection
- (iv) data analysis
- (v) findings and conclusions
- (vi) validity
- (vii) relevance

The coding forms for data extraction are included in Appendix 2.

Data synthesis:

The qualitative findings and conclusions of each study will be copied verbatim as reported by study authors into NVivo from QSR Software, a specialist software package for undertaking qualitative analysis of textual data. Examining these findings, every sentence or paragraph within the report of findings will be assigned a code to describe it. Two independent reviewers will look for similarities and differences between the codes in order to group them into a hierarchical tree structure of descriptive themes. A narrative summary of the findings across the studies will then be written and organized by the descriptive themes.

## References

Critical Appraisal Skills Programme (CASP), [www.phru.nhs.uk/casp/casp.htm](http://www.phru.nhs.uk/casp/casp.htm)

GREENHALGH, T. (2001) *How to Read a Paper*, 2nd ed., BMJ Books.

KITCHENHAM, B.A. (2004) *Procedures for Performing Systematic Reviews*, Keele University, Technical Report TR/SE-0401-ISSN:1353-7776 and NICTA Technical Report 0400011T.1.

KITCHENHAM, B.A., PFLEEGER, S.L., PICKARD, L.M., JONES, P.W., HOAGLIN, D.C., EL EMAM, K., AND ROSENBERG, J. (2002) Preliminary Guidelines for Empirical Research in Software Engineering, *IEEE Transactions on Software Engineering*, 28(8):721-734.

## Appendix 1: Quality assessment form

### Screening questions

<p><b>1. Is this a research paper?</b>            Consider:            – Is the paper based on research (or is it merely a “lessons learned” report based on expert opinion)</p>	<p><input type="checkbox"/> Yes      <input type="checkbox"/> No</p>
<p><b>2. Is there a clear statement of the aims of the research?</b>            Consider:            – Is there a rationale for why the study was undertaken?            – Is the study’s focus or main focus on Agile Software Development            – Does the study present empirical data?            – Is there a clear statement of the study’s primary outcome (i.e. time-to-market, cost, or product or process quality)?</p>	<p><input type="checkbox"/> Yes      <input type="checkbox"/> No</p>
<p><b>3. Is there an adequate description of the context in which the research was carried out</b>            Consider if the researcher has identified:            – The industry in which products are used (e.g. banking, telecommunications, consumer goods, travel, etc)?            – The nature of the software development organization (e.g. in-house department or independent software supplier)?            – The skills and experience of software staff (e.g. with a language, a method, a tool, an application domain)?            – The type of software products used (e.g. a design tool, a compiler)            – The software processes being used (e.g. a company standard process, the quality assurance procedures, the configuration management process)?</p>	<p><input type="checkbox"/> Yes      <input type="checkbox"/> No</p>

**If questions 1 or 2 receive a “No” response do not continue with the quality assessment.**

## Detailed questions

<p><b>Research design</b></p> <p>4. Was the research design appropriate to address the aims of the research?  <i>Consider:</i>          – Has the researcher justified the research design (e.g. have they discussed how they decided which methods to use)?</p>	<p><input type="checkbox"/> Yes      <input type="checkbox"/> No</p>
<p><b>Sampling</b></p> <p>5. Was the recruitment strategy appropriate to the aims of the research?  <i>Consider:</i>          – Has the researcher explained how the participants or cases were identified and selected?          – Are the cases defined and described precisely?          – Was the cases representative of a defined population?          – Have the researchers explained why the participants or cases they selected were the most appropriate to provide access to the type of knowledge sought by the study?          – Was the sample size sufficiently large?</p>	<p><input type="checkbox"/> Yes      <input type="checkbox"/> No</p>
<p><b>Control group</b></p> <p>6. Was there a control group with which to compare treatments?  <i>Consider:</i>          – How were the controls selected?          – Were they representative of a defined population?          – Was there something special about the controls?          – Was the non-response high? Could non-respondents be different in any way?</p>	<p><input type="checkbox"/> Yes      <input type="checkbox"/> No</p>
<p><b>Data collection</b></p> <p>7. Was the data collected in a way that addressed the research issue?  <i>Consider:</i>          – Were all measures clearly defined (e.g. unit and counting rules)?          – Is it clear how data was collected (e.g. semi-structured interviews, focus group etc.)?          – Has the researcher justified the methods that were chosen?          – Has the researcher made the methods explicit (e.g. is there an indication of how interviews were conducted, did they use an interview guide)?          – If the methods were modified during the study has the researcher explained how and why?          – If the form of data is clear (e.g. tape recording, video material, notes etc.)?          – If quality control methods were used to ensure completeness and accuracy of data collection?</p>	<p><input type="checkbox"/> Yes      <input type="checkbox"/> No</p>



<p><b>Data analysis</b></p> <p>8. Was the data analysis sufficiently rigorous?  <i>Consider:</i></p> <ul style="list-style-type: none"> <li>– Was there an in-depth description of the analysis process?</li> <li>– If thematic analysis was used is it clear how the categories/ themes were derived from the data?</li> <li>– Has sufficient data been presented to support the findings?</li> <li>– To what extent has contradictory data been taken into account?</li> <li>– If quality control methods were used to verify the results?</li> </ul>	<p><input type="checkbox"/> Yes      <input type="checkbox"/> No</p>
<p><b>Reflexivity (research partnership relations/recognition of researcher bias)</b></p> <p>9. Has the relationship between researcher and participants been adequately considered?  <i>Consider:</i></p> <ul style="list-style-type: none"> <li>– Has the researcher critically examined their own role, potential bias and influence during formulation of research questions, sample recruitment, data collection, and analysis and selection of data for presentation?</li> <li>– How the researcher responded to events during the study and whether they considered the implications of any changes in the research design?</li> </ul>	<p><input type="checkbox"/> Yes      <input type="checkbox"/> No</p>
<p><b>Findings</b></p> <p>10. Is there a clear statement of findings?  <i>Consider:</i></p> <ul style="list-style-type: none"> <li>– Are the findings explicit (e.g. magnitude of effect)?</li> <li>– Has an adequate discussion of the evidence, both for and against the researcher's arguments, been demonstrated?</li> <li>– Has the researcher discussed the credibility of their findings (e.g. triangulation, respondent validation, more than one analyst)?</li> <li>– Are limitations of the study explicitly discussed?</li> <li>– Are the findings discussed in relation to the original research questions?</li> <li>– Are the conclusions justified from the results?</li> </ul>	<p><input type="checkbox"/> Yes      <input type="checkbox"/> No</p>
<p><b>Value of the research</b></p> <p>11. Is the study of value for research or practice?  <i>Consider:</i></p> <ul style="list-style-type: none"> <li>– Does the researcher discuss the contribution the study makes to existing knowledge or understanding (e.g. do they consider the findings in relation to current practice or relevant research-based literature)?</li> <li>– Does the research identify new areas where research is necessary?</li> <li>– Does the researcher discuss whether or how the findings can be transferred to other populations or considered other ways the research can be used?</li> </ul>	<p><input type="checkbox"/> Yes      <input type="checkbox"/> No</p>

## Appendix 2: Data extraction form

<p><b>Study description</b></p> <ol style="list-style-type: none"> <li>1. Study identifier:</li> <li>2. Date of data extraction</li> <li>3. Bibliographic reference:</li> <li>4. Type of article:</li> <li>5. Study aims:</li> <li>6. Objectives:</li> <li>7. Design of study:</li> <li>8. Research hypothesis:</li> <li>9. Definition of agile software development given in study:</li> <li>10. Sample description</li> <li>11. Setting of study</li> <li>12. Control group</li> <li>13. Data collection</li> <li>14. Data analysis</li> </ol>	<p>Unique id for the study</p> <p>Author, year, title, source</p> <p>Journal article, conference paper, workshop paper, book section</p> <p>What were the aims of the study?</p> <p>What were the objectives?</p> <p>Qualitative, quantitative (experiment, survey, case study, action research)</p> <p>Statement of hypotheses, if any</p> <p>Verbatim from the study</p> <p>Size, students, professionals (age, education, experience)</p> <p>Industry, in-house/supplier, products and processes used</p> <p>Yes, no (#groups, sample size)</p> <p>How was the data obtained? (questionnaires, interviews, forms)</p> <p>How was the data analyzed? (qualitative, quantitative)</p>
<p><b>Study findings</b></p> <ol style="list-style-type: none"> <li>1. Findings and conclusions</li> <li>2. Validity</li> <li>3. Relevance</li> </ol>	<p>What were the findings and conclusions? (verbatim from the study)</p> <p>Limitations, threats to validity</p> <p>Research, practice</p>