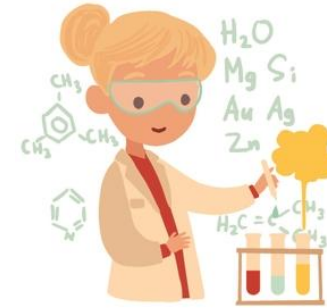


Revisiones Sistemáticas de Literatura

Ingeniería de Software Basada en Evidencias

Agregación de evidencia

Resultado de investigación (evidencia).



Resultado de muchas investigaciones (agregación de evidencias).



Práctica basada en evidencias

- Existe la necesidad de agregar evidencia desde múltiples estudios empíricos.
- El proceso de recolección y síntesis debe contar con cierto estándar científico y rigurosidad metodológica.
- La práctica basada en evidencias (EBP) surge en medicina y luego se propaga a otras disciplinas.



Cochrane
Community

Trusted evidence.
Informed decisions.
Better health.



Ing. de soft. basada en evidencias (EBSE)

- Se ocupa de determinar qué funciona, cuándo y dónde, en términos de prácticas de ingeniería de software, herramientas y estándares.
- Su técnica principal es la Revisión Sistemática de Literatura (SLR).
- Una SLR es planificada formalmente y ejecutada de manera sistemática y metódica.



Pasos de EBSE

1. Convertir el problema en una pregunta que se pueda responder.
2. Buscar en la literatura la mejor evidencia disponible para su respuesta.
3. Evaluar críticamente la evidencia en su:
 - validez – qué tan cerca está de la verdad,
 - impacto – el “tamaño” de los efectos observados, y
 - aplicabilidad – qué tan útil es.
4. Integrar la evidencia con la experiencia práctica y los valores del cliente.
5. Evaluar la eficacia y eficiencia de los pasos previos para mejorarlos.

Ejemplo de problemas abordados con EBSE

- ¿Cuál es el costo de desarrollar software usando cascada o sus variaciones versus usando desarrollo iterativo e incremental?
- ¿Deberíamos esperar estimaciones de esfuerzo más precisas al aplicar el juicio de expertos o modelos paramétricos?
- ¿Qué motiva a los ingenieros de software a ser más productivos?
- ¿Cuáles son algunas estrategias para mejorar el involucramiento de los clientes en proyectos de desarrollo de software que utilizan metodologías ágiles?

Métodos dentro de EBSE

- Revisión Sistemática de Literatura (SLR).
- Estudio de Mapeo Sistemático – su propósito es encontrar y clasificar estudios primarios (es más exploratorio).
- Estudio Terciario – revisión de revisiones.
- Revisión Rápida (rapid review) – revisión limitada, por tiempo, por alcance o por actividades.
- Revisiones Multivocales de Literatura - incluyen literatura gris (reportes técnicos, tesis y hasta blogs).

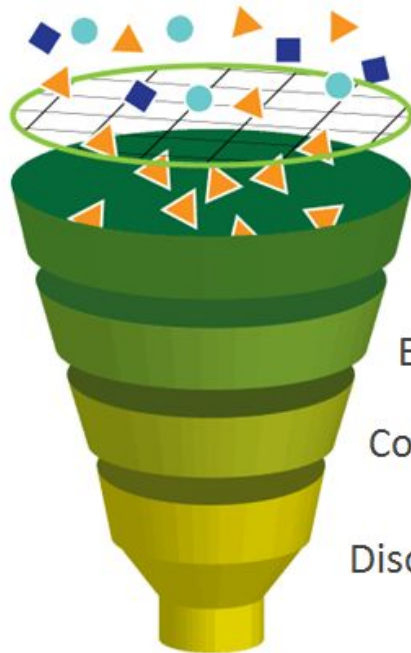
Proceso de revisión

By Centre for Health Communication and Participation La Trobe University,
Australasian Cochrane Centre [CC BY-SA 3.0
(<http://creativecommons.org/licenses/by-sa/3.0>)], via Wikimedia Commons

Identificar problema y preguntas

Elaborar un plan (protocolo)

Buscar estudios



Filtrar y seleccionar estudios

Extraer datos de los estudios

Evaluar la calidad de los estudios

Combinar los datos (síntesis)

Discutir y concluir resultados generales

Revisión sistemática

Difusión

Revisión sistemática de literatura (SLR)

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)*

Planificar la Revisión

- Identificar la necesidad de una revisión
- Especificar las preguntas de investigación
- Desarrollar el protocolo de revisión

Planificar la Revisión

- Identificar la necesidad de una revisión
 - Chequear si no existen revisiones previas o en curso y si una nueva revisión es necesaria.
 - Si hay SLRs que contestan la pregunta de interés, entonces se deben evaluar si cuentan con calidad suficiente.
- Especificar las preguntas de investigación
 - Buscamos preguntas interesantes para investigadores como para profesionales, que darán lugar a cambios en las prácticas actuales de la ingeniería de software o que mostrará discrepancias entre las creencias comúnmente aceptadas y la realidad.
- Desarrollar el protocolo de revisión
 - Especifica los métodos que se utilizarán y su definición de antemano tiene como objetivo minimizar sesgos.

Realizar la Revisión

- Identificar la investigación relevante
- Seleccionar los estudios primarios
- Evaluar la calidad de los estudios

Realizar la Revisión

- Identificar la investigación relevante

- Se utiliza una estrategia de búsqueda (definida en el protocolo) para obtener una lista de todas las publicaciones relevantes para las preguntas de investigación.
- 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.
- Es posible hacer búsquedas automáticas o manuales.

- Seleccionar los estudios primarios

- Una vez que se han obtenido los estudios primarios potenciales es necesario evaluar su relevancia. Para esto se utilizan criterios de inclusión y exclusión.

- Evaluar la calidad de los estudios

- Busca asegurar que los hallazgos de un estudio son relevantes y no sesgados. Instrumentos de calidad -> evaluaciones detalladas, en gral. checklists.

Realizar la Revisión (2)

- Extraer Datos
- Síntesis de Datos
- Informar la revisión

Realizar la Revisión (2)

• Extraer Datos

- En esta etapa se recolecta la información según los formularios definidos en el protocolo. Debería incluir las siguientes partes:
 - información de la extracción (responsable y fecha de la extracción),
 - información general de los estudios - id y detalles de publicación,
 - preguntas para responder las preguntas de la revisión,
 - preguntas para evaluar la calidad de los estudios

• Síntesis de Datos

- Responder las preguntas de investigación mediante integrar los resultados incluidos en los estudios primarios. Puede ser síntesis cualitativa o cuantitativa.
- Comenzar con un resumen de los estudios encontrados - una tabla con detalles importantes como son: tipo, intervenciones, detalle de los participantes, resultados, etc. Incluir elementos de la calidad de los estudios.

• Informar la revisión

- Redactar los resultados de la revisión y difundir los resultados a potenciales interesados.

Estudios de Mapeo

- Son un tipo de SLR, aunque su propósito es encontrar y clasificar estudios primarios dentro de un tópico específico.
- La diferencia principal
 - Una SLR intenta agregar los estudios primarios en términos de los resultados de la investigación e investiga si son consistentes o contradictorios.
 - Un Mapeo aspira solamente a clasificar la literatura relevante y a clasificar los estudios con respecto a categorías definidas

Estudios de Mapeo

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).

Estudios de Mapeo

- Esquemas de clasificación

- Presentación de resultados

Estudios de Mapeo

- Esquemas de clasificación

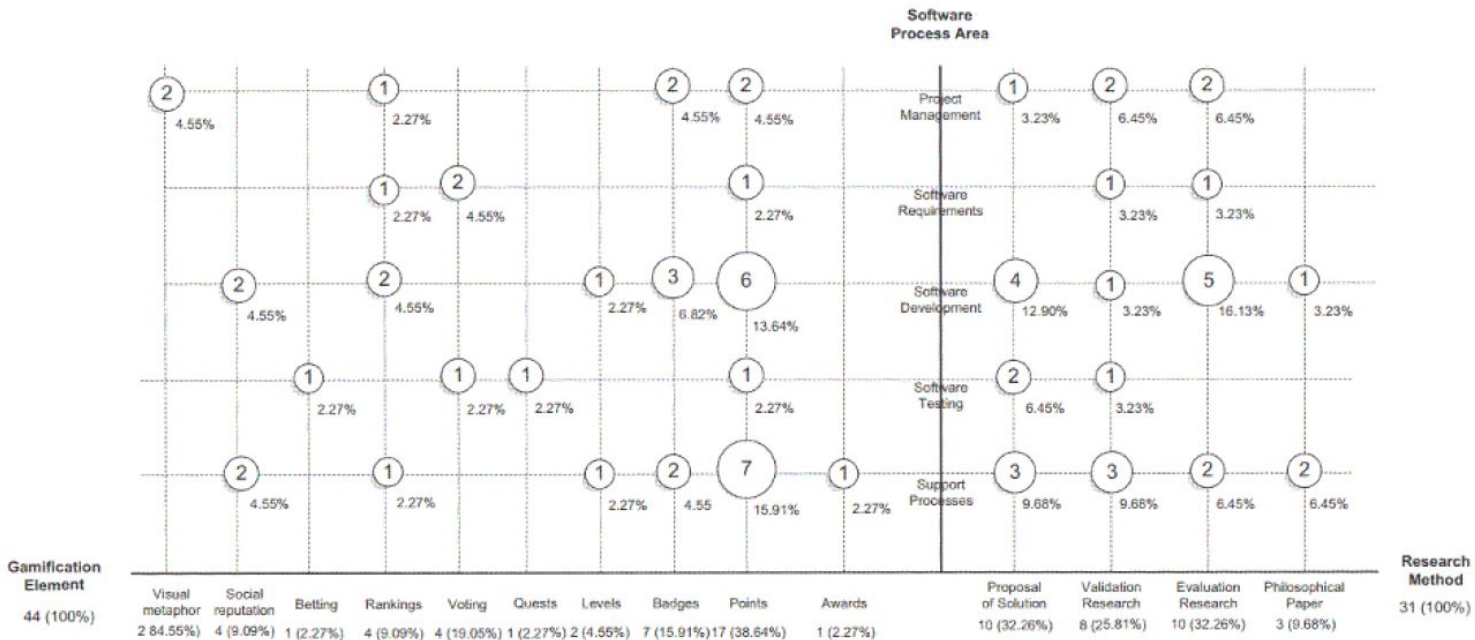
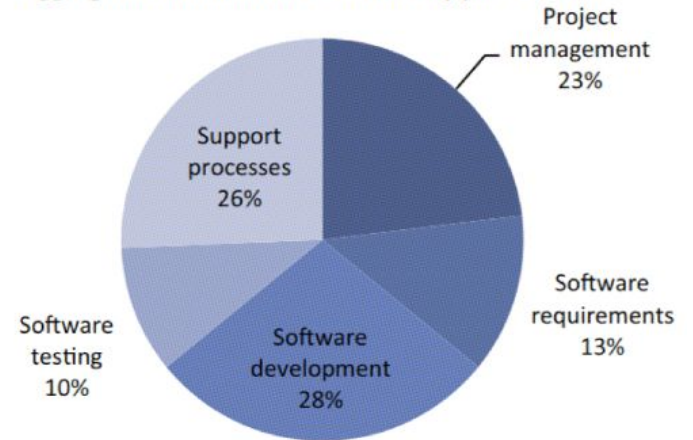
- En general se trabaja con uno o más esquemas para categorizar los estudios encontrados, por ejemplo, identificando a qué país pertenece o estableciendo de qué tópico trata a partir de una lista de temas.
- Para elegir los esquemas se puede consultar investigación previa o bibliografía relevante antes de comenzar.

- Presentación de resultados

- Como buscamos clasificar los estudios primarios es frecuente incluir gráficos para mostrar la distribución de los estudios según una dimensión (una categoría, se pueden utilizar gráficos de torta u otros) o varias (cruzando dos o tres categorías, por ejemplo en un gráfico de burbujas).

Estudios de Mapeo

Aggregated distribution of studies by process area



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.

Rapid Reviews

- Tienen como propósito entregar evidencia con tiempos razonables.
- Para esto omiten o simplifican ciertos pasos.
- Ejemplo: Limitan literatura (1 motor o biblioteca), una persona sola hace la selección, no se hace evaluación de calidad, presentan los resultados sin una síntesis formal.
- Se puede además buscar feedback temprano en “clientes”.
- Limitaciones: no son tan confiables como la SLRs (cantidad de evidencia, calidad, sesgo, etc.). Sirven para preguntas exploratorias o para tomar decisiones rápidas que no necesitan tanta evidencias (por ejemplo, elegir posibles estrategias para involucrar más a los clientes en un proyecto ágil).

Evidence Briefings

- Un medio que permite presentar los resultados de un estudio secundario en una sola página que resulten de interés para practicantes de la ingeniería de software.

6 Empirical Software Engineering Group

Centro de Informática

UNIVERSIDADE FEDERAL DE PERNAMBUCO

1

HARMFULNESS OF CODE DUPLICATION

2 This briefing reports evidence on circumstances under which code duplication harms system quality based on scientific evidence from a systematic review.

FINDINGS

Figure 1 contains a causal model of the effects of duplication on system quality. The model has been constructed after analysis of the reviewed literature. The names in the boxes are variables and the arrows are causal effects. Each arrow represents a separate hypothesis such as, "duplication increases cochange". This should be read that: in general, all else being equal, a system with more duplication will exhibit more cochange than an equivalent system with less duplication.

3

Co-Change:

- A significant portion of clones are co-changed. For cochanging clones that cannot be refactored, consider using change propagating tools, such as Linked Editing.
- Of the clones that are removed in later versions, a minority is removed because of different changes, while the rest is presumably removed through refactoring.
- Clones that are not refactorable tend to live longer and are very likely to exhibit cochange. 45% to 55% of clone sets are cochanged during a long part of the life cycle of systems.
- Clones with class-level granularity have more cochange than clones with smaller granularities. These also cause more increase in code size. These should be the first to refactor.
- Files between which there is a very high

ratio of cloning are very often changed in the same release.

INCONSISTENT CHANGE

- A small portion of clones are changed inconsistently which is repaired in later versions. However only a part of the cases in which a change is applied to other occurrences later on (late propagation) stem from inconsistent change.
- In a small number of cases it was observed that the act of copying code led to errors. Consider using the approaches mentioned as errors in the code, especially when a lot of copy/paste has happened.

CODE COMPREHENSION

- Programmers experience that duplication hinders code comprehension. Consider adding close definition results to system documentation to aid comprehension.
- Issues with code ownership may introduce duplications.
- Changes that involve aspects [in the ADP sense] are especially hard to analyze. When aspect-related code is cloned and changes to those aspects are expected in the future, use ADP techniques to avoid cochange.
- Duplicated code is not necessarily more complex or harder to understand than the equivalent non-duplicated code. Be careful when refactoring clones to not increase the complexity.

Code Size

- It is simple to measure the total size of code involved in a clone set. However it is difficult to estimate the size of code needed in an alternative solution, for example after refactoring. Removal of clones does not always result in reducing the code size. Be careful when refactoring clones to not increase the code size.
- Researchers should be aware of types of duplication that go beyond syntactical clones.
- Clones created by different programmers are harder to understand and harder to refactor than clones created by one programmer.

4 Words:

Code duplication
Code clone
Software Maintenance

Who is this briefing for?

Software engineers/practitioners who want to make decisions about code duplication based on scientific evidence.

Where the findings come from?

All findings of this briefing were extracted from the systematic review conducted by Horájk et al.

What is a systematic review?

cin.ufpe.br/insig/sirs

What is included in this briefing?

The main findings of the original systematic review.

What is not included in this briefing?

Additional information not presented in the original systematic review.

To access other evidence briefings on software engineering:

cin.ufpe.br/insig/briefings

For additional information about ESEG:

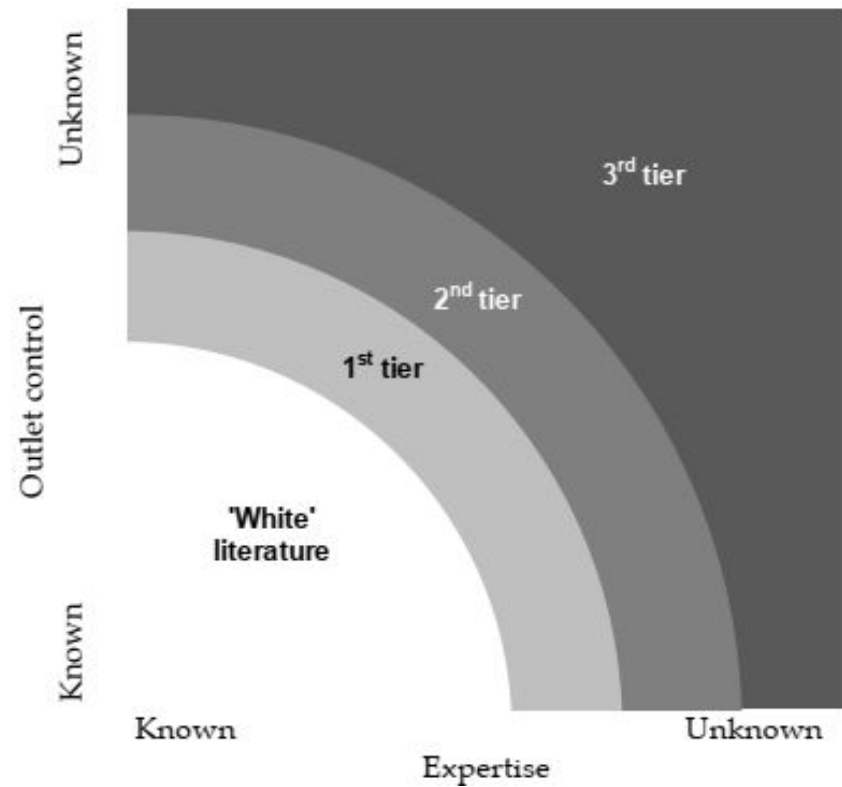
cin.ufpe.br/insig

5 ORIGINAL SYSTEMATIC REVIEW REFERENCE

Wierke Horájk, Maria Laura Pinheiro, and Rolf Wherrings. 2009. Harmfulness of code duplication: a structured review of the evidence. In Proceedings of the 13th International Conference on Evaluation of Assessment in Software Engineering (EASE'09), David Butler, Mark Turner, and Mahmood Niaz (Eds.), British Computer Society, Swinton, UK, UK, 489-7.

Revisiones Multivocales de Literatura

Se componen de todos los escritos accesibles sobre un tema común, incorporando las visiones de diversos grupos de autores (académicos, profesionales, periodistas, centros de políticas, oficinas estatales de educación, distritos escolares locales, empresas independientes de investigación y desarrollo, y otros). Los escritos aparecen en una variedad de formas, reflejan diferentes propósitos, perspectivas y bases de información, y abordan diferentes aspectos del tema e incorporan diferentes lógicas de investigación o no investigación



3rd tier GL:

Low outlet control/ Low credibility:

Blogs, emails, tweets

2nd tier GL:

Moderate outlet control/ Moderate credibility:

Annual reports, news articles, presentations, videos, Q/A sites (such as StackOverflow), Wiki articles

1st tier GL:

High outlet control/ High credibility:

Books, magazines, government reports, white papers