



Universidade Federal de Pernambuco
Centro de Ciências Exatas e Natureza
Centro de Informática

Pós-Graduação em Ciência da Computação

**UM MAPEAMENTO SISTEMÁTICO DE
MECANISMOS PARA GUIAR ESTUDOS
EMPÍRICOS EM ENGENHARIA DE
SOFTWARE.**

Adauto Trigueiro de Almeida Filho

DISSERTAÇÃO DE MESTRADO

Recife
Fevereiro de 2011

Universidade Federal de Pernambuco
Centro de Ciências Exatas e Natureza
Centro de Informática

Adauto Trigueiro de Almeida Filho

**UM MAPEAMENTO SISTEMÁTICO DE MECANISMOS PARA
GUIAR ESTUDOS EMPÍRICOS EM ENGENHARIA DE
SOFTWARE.**

*Trabalho apresentado ao Programa de Pós-Graduação em
Ciência da Computação do Centro de Informática da Uni-
versidade Federal de Pernambuco como requisito parcial
para obtenção do grau de Mestre em Ciência da Com-
putação.*

Orientador: *Sérgio Castelo Branco Soares*

Recife
Fevereiro de 2011

AGRADECIMENTOS

Aos meus pais, que me apoiaram durante o mestrado.

A minha noiva Larissa Coutinho, pela paciência e cumplicidade.

Aos meus futuros sogro e sogra, pelo apoio.

A todos meus parentes, que me ajudaram durante esta etapa.

Ao meu orientador Sérgio Soares, obrigado pela oportunidade e pelo empenho durante o mestrado.

A FACEPE, que acreditou no projeto e nos financiou, muito obrigado.

Obrigado a Emanuel Barreiros, Helaine Lins, Juliana Gonçalves e a todos os amigos que conviveram comigo durante o mestrado.

Coming together is a beginning. Keeping together is progress. Working together is success.

—HENRY FORD

RESUMO

Estudos empíricos têm se mostrado mecanismos relevantes para o avanço científico em diversas áreas, como por exemplo, as Ciências Sociais e a Medicina. Tais áreas possuem considerável experiência na condução de estudos dessa natureza. Em Engenharia de Software (ES) não é diferente, pois desde a década de 1980 estudos empíricos têm sido evidenciados como um importante instrumento para o desenvolvimento da área, que é fortemente dependente de atividades humanas. Nesse cenário, é essencial ter conhecimento para conduzir os principais tipos de estudos empíricos encontrados em ES, a saber: experimento controlado, quasi-experimento, *survey*, etnografia, pesquisa-ação e estudo de caso. A partir desse contexto, esta pesquisa visa selecionar os trabalhos mais relevantes da área que foram capazes de desenvolver algum tipo de mecanismo específico para ES que servisse de guia para os estudos empíricos citados anteriormente. Essa seleção foi realizada através de um mapeamento sistemático de estudos, o qual utilizou buscas automatizadas com a ajuda de cinco engines de busca de bibliotecas digitais, além de busca manual em periódicos e conferências relevantes para a área. No total, foram encontrados 7101 estudos, dos quais 23 foram identificados como trabalhos primários relevantes e serviram como fonte para coleta de evidências para responder as três perguntas de pesquisa propostas. Após a análise, conclui-se que há alguns esforços em fornecer guias específicos que apoiem a condução de estudos empíricos em ES, inclusive, aproveitando experiências de outras áreas. No entanto, ainda há lacunas na área de ES empírica a serem preenchidas no que diz respeito à disponibilidade de material guia específico para ES, como por exemplo, em etnografia e pesquisa-ação, além da padronização dos procedimentos. O preenchimento de tais lacunas é essencial para fomentar a condução de tais pesquisas empíricas que são fundamentais para o desenvolvimento da ES.

Palavras-chave: Engenharia de Software, Engenharia de Software Empírica, Engenharia de Software Experimental, Experimento, Quasi-experimento, Estudo de Caso, Etnografia, Pesquisa Ação, Survey.

ABSTRACT

Empirical studies have been reported as relevant mechanisms to the scientific advance in a number of areas, such as, the Social Sciences and Medicine. Such fields have considerable experience in conducting studies of this nature. In Software Engineering (SE) it is not different, because since the 80s empirical studies have been evidenced as an important instrument to the development of the area, which is strongly dependent on human activities. In this scenario, it is essential to have knowledge to conduct the main types of empirical studies found in SE, namely: controlled experiment, quasi-experiment, survey, ethnography, action-research and case study. In this context, the this research aims to select the most relevant studies that were able to develop some kind of mechanism, specific to SE, that can be used to guide the categories of empirical studies cited before. The studies selection was performed through a systematic mapping study, which used automated search engines from five digital libraries and manual searches on relevant journals and conferences. In total, 7101 studies have been found, from which 23 were identified as relevant primary studies and served as sources to collect evidences to answer the 3 proposed research questions. After the analysis, the study concludes that there are efforts to provide specific guides to support the conduction of empirical studies in SE, including experiences from other areas. However, there are still gaps to be filled in SE regarding the availability of material specific to SE, such as ethnography, action-research, and procedure standardization. Filling those gaps is crucial to foster the conduction of empirical research, which is fundamental to the development of SE.

Keywords: Software Engineering, Empirical Software Engineering, Experimental Software Engineering, Experiment, Quasi-experiment, Casy Study, ethnography, Action Research, Survey.

SUMÁRIO

Capítulo 1—Introdução	1
1.1 Motivação	1
1.2 Objetivos	3
1.2.1 Objetivos Gerais	3
1.2.2 Objetivos Específicos	3
1.3 Estrutura da Dissertação	3
Capítulo 2—Fundamentação Teórica	5
2.1 Engenharia de Software	5
2.2 Estudos Empíricos em Engenharia de Software	6
2.3 Tipos de Estudos Empíricos em Engenharia de Software	8
2.4 Engenharia de Software Baseada em Evidências	9
2.5 Resumo	14
Capítulo 3—Metodologia	15
3.1 Classificação da Pesquisa	15
3.1.1 Classificação segundo Cooper	16
3.2 Ciclo da Pesquisa	18
3.2.1 Mapeamento Sistemático	19
3.2.2 Escopo e Questões de Pesquisa	20
3.2.3 Estratégia de Busca	21
3.2.4 Processo de Busca	22
3.2.5 Critérios de Inclusão/Exclusão e Procedimentos de Seleção	23
3.2.6 Extração dos dados	23
3.2.7 Estratégia de Síntese dos Dados	24
3.3 Resumo	25

Capítulo 4—Resultados	26
4.1 Extração e análise dos dados	26
4.2 Mapeamento das evidências	31
4.2.1 Tipo de Mecanismo	32
4.2.2 Tipo de Estudos Empíricos	40
4.2.3 Estrutura dos Mecanismos	45
4.3 Discussão sobre os Resultados	53
4.4 Resumo	57
Capítulo 5—Considerações Finais	58
5.1 Ameaças a validade	58
5.2 Trabalhos futuros	59
5.3 Conclusões	59
Apêndice A—Estudos Primários	61
Apêndice B—Estudos Excluídos	63
Apêndice C—Protocolo do Mapeamento Sistemático	74
C.1 Equipe	74
C.2 Introdução	74
C.3 Escopo e Questões de Pesquisa	76
C.4 Estratégia e Fontes de Busca	77
C.5 Seleção de Estudos	78
C.5.1 Processo Modelo de Seleção	78
C.5.2 Critérios de Inclusão/Exclusão	79
C.5.3 Etapas de Seleção dos Estudos	79
C.6 Extração dos dados	81
C.6.1 Formulário A	81
C.6.2 Formulário B	81
C.6.3 Formulário C	82
C.7 Estratégia de Síntese dos Dados	82

LISTA DE FIGURAS

3.1	Etapas da Pesquisa	19
4.1	Resultado da busca automatizada.	27
4.2	Resultado da busca manual.	28
4.3	Resultado das buscas automatizada e manual juntas.	28
4.4	Resultado seleção de trabalhos potencialmente relevantes para o estudo.	29
4.5	Resultado seleção de trabalhos primários que compõem o estudo.	30
4.6	Distribuição dos estudos ao longo dos anos.	31
4.7	Distribuição dos estudos pelos países das instituições dos pesquisadores.	32
4.8	33
4.9	Distribuição dos EP selecionados por tipo de estudo abordado.	41
4.10	Experimentation Process [1].	46
4.11	Activities in the process to perform on-line surveys [2].	47
4.12	Sumário do framework para experimentação [3].	50
4.13	Visão geral do processo experimental [4].	50
4.14	O paradigma Goal/Question/Metric (GQM). Goals= G_i , Questions= Q_i e Metrics= M_i [5].	52
4.15	Passos do protocolo baseado em diálogo [6].	53
4.16	Estrutura base compartilhada entre maior parte dos guias.	56
C.1	Processo modelo usado para seleção de estudos primários	79
C.2	Primeira Etapa da Seleção	80
C.3	Segunda Etapa da Seleção	81

LISTA DE TABELAS

3.1	Classificação da Pesquisa.	15
3.2	Classificação segundo Cooper.	17
4.1	Evolução do procesos de seleção.	30
4.2	Resumo dos estudos por tipo de mecanismo encontrado.	33
4.3	Resumo dos resultados extraídos para questões Q1 e Q2	55

1.1 MOTIVAÇÃO

Ao longo dos últimos anos é observado na Engenharia de Software (ES) um crescente interesse em conduzir estudos empíricos para avaliar o uso de novas tecnologias, métodos, técnicas, ferramentas, linguagens, etc. Tais estudos têm ganhado importância devido ao apoio à tomada de decisão fornecida aos profissionais e na diminuição do prazo de transição da tecnologia da academia para indústria. Isso ocorre devido à geração de evidências sobre eficiência e eficácia das tecnologias que passam por avaliações sistemáticas, gerando assim uma base de conhecimento que pode ser utilizada para decidir o que deve ser usado ou não no decorrer do desenvolvimento de soluções em engenharia de software.

Um dos primeiros estudos que evidencia a importância dessa prática foi publicado em 1986 por Basili et al. [3]. Nos últimos 25 anos, outros estudos têm reforçado a importância dos estudos empíricos em engenharia de software, por muitas vezes, referenciada como Engenharia de Software Empírica (ESE). Tais estudos propuseram ambientes de apoio, *guidelines*, metodologias, *frameworks* e outros mecanismos que permitem projetar, planejar, executar, analisar e empacotar estudos empíricos em ES [7, 8, 9, 10, 11, 12, 13, 14].

De acordo com Pfleeger [15], a condução de estudos experimentais ajudaria a entender as chances que, sob certas condições, uma tecnologia conduzirá o aprimoramento de alguma característica do software. Próximo a esse pensamento, Basili et al. [16] enfatizam em seu estudo que o progresso de qualquer disciplina científica, envolve a construção de modelos que possam ser testados através de estudos empíricos, fazendo com que eles possam fornecer uma forma de testar, refinar e refutar o entendimento de um fenômeno.

O desenvolvimento de soluções que envolvem sistemas de software é intensivo em atividades humanas, e, diferente de outros meios de produção, os desenvolvedores de software não repetem completamente a produção de um mesmo produto, dessa forma cada solução de software desenvolvida é diferente da anterior [17]. Portanto, a interação humana é um fator importante na ES, o que a torna próxima, nesse sentido, a ciências sociais e comportamentais, que já possuem experiência na prática de estudos empíricos

em seus campos de pesquisa [18]. Nesse sentido, estudos empíricos podem auxiliar no entendimento e identificação da influência da interferência humana no decorrer do desenvolvimento de sistemas. Segundo Amaral [19], um exemplo de um estudo empírico como a experimentação oferece um modo sistemático, disciplinado, computável e controlado para avaliação de novas tecnologias, já que a intuição, opinião e especulação não podem ser consideradas fontes confiáveis de conhecimento.

Portanto, novos métodos, técnicas, processos, linguagens e ferramentas não deveriam ser apenas sugeridas, publicadas e comercializadas. Antes disso deveriam ser rigorosamente comparadas com outras já existentes, conhecendo assim os riscos inerentes e justificando a sua adoção. Visando atingir tais objetivos para a pesquisa empírica em ES, existem iniciativas que visam fomentar a execução de tais estudos. Um exemplo é o uso de *benchmarks* [20, 21], que pode ser definido como uma maneira de encapsular as informações essenciais e procedimentos (hipóteses, tratamentos, variáveis independentes, objeto de controle, objeto do estudo, resultados e etc) utilizados na execução de estudos empíricos, podendo ser usado para responder diversas questões de forma objetiva [22]. *Benchmarks* são um referencial para realizar comparações e análises objetivas. Partindo desta idéia, Moura [23] propôs a criação de um *benchmark framework* (BF) com foco em avaliação de manutenibilidade de software orientado a aspectos. O autor do trabalho define um BF como:

Conjuntos de atributos capazes de identificar características específicas de diferentes aplicações (*framework*), de modo que seja possível avaliá-las e compará-las (*benchmarking*) sob um determinado contexto (manutenibilidade de software OA).

O *benchmark framework* tem a intenção de apoiar pesquisadores e profissionais em duas tarefas: (1) no desenvolvimento, adaptação e avaliação de aplicações *benchmark* para manutenção de software orientado a aspectos e (2) na elaboração, avaliação e empacotamento de estudos empíricos sobre manutenibilidade de softwares orientado a aspectos.

Contudo, o trabalho de Moura [23] não define de forma clara o processo no tocante a elaboração, execução, avaliação e empacotamento de estudos empíricos, limitando-se a uma superficial descrição do mesmo. Com a finalidade de preencher esta lacuna, esta dissertação tem como foco principal agregar o conhecimento sobre a condução de estudos empíricos em ES, que segundo Easterbrook [13] são: experimento controlado, quasi-experimento, estudo de caso, *survey*, pesquisa-ação e etnografia. E dessa forma, contribuir para a continuidade do estudo realizado por Moura [23] e paralelamente para o fomento da prática de estudos empíricos em engenharia de software.

1.2 OBJETIVOS

1.2.1 Objetivos Gerais

Este trabalho tem como objetivo agregar o conhecimento disponível na literatura sobre mecanismos para guiar estudos empíricos. A partir do conjunto de mecanismos encontrados, os que forem adequados farão parte do processo de utilização do *benchmark framework* definido por [23].

1.2.2 Objetivos Específicos

Com o resultado proveniente do mapeamento sistemático realizado pretende-se alcançar os seguintes objetivos específicos:

- Auxiliar a criação de planos e execução de estudos empíricos em ES;
- Facilitar a geração de evidências empíricas em ES;
- Favorecer a popularização de técnicas empíricas em pesquisas de Engenharia de Software.

1.3 ESTRUTURA DA DISSERTAÇÃO

O estudo está estruturado da seguinte forma:

- No Capítulo 2 é apresentado o referencial teórico, que contem a revisão de conceitos que fundamentam esta dissertação.
- No Capítulo 3 é descrita a metodologia que foi adotada para realizar esta pesquisa. É apresentada a classificação da pesquisa com a apresentação de um quadro metodológico, as principais etapas deste estudo, os procedimentos seguidos pelo mapeamento sistemático, com apresentação do protocolo definido, a forma como os dados foram extraídos, analisados e sintetizados.
- No Capítulo 4 são apresentados os resultado do mapeamento sistemático. No início são apresentadas informações gerais sobre o processo de pesquisa e seleção dos estudos primários que participaram do estudo, como as principais fontes, número de estudos retornados, distribuições por países e temporal, e outras informações. Ao final do capítulo são apresentadas as evidências encontradas que serviram para responder as perguntas de pesquisa e o mapeamento resultante.

- No Capítulo 5, o capítulo final, são apresentadas as limitações e ameaças a validade do trabalho, propõe trabalhos futuros a partir dos resultados e finaliza com as conclusões.

CAPÍTULO 2

FUNDAMENTAÇÃO TEÓRICA

Neste capítulo será apresentada a fundamentação teórica obtida através de uma revisão bibliográfica tradicional (informal). Os tópicos que serão abordados são: engenharia de software, estudo empíricos em engenharia de software, os principais tipos de estudos realizados em engenharia de software e a engenharia de software baseada em evidências. As subseções a seguir mostram em detalhes cada tópico citado.

2.1 ENGENHARIA DE SOFTWARE

Algumas definições para engenharia de software encontrada na literatura são:

- O estabelecimento e uso de sólidos princípios de engenharia para obter software confiável e que trabalhe de forma eficiente em máquinas reais, Fritz Bauer em 1969 [24].
- A aplicação de uma abordagem sistemática, disciplinada e quantificável para o desenvolvimento, operação e manutenção de software. O estudo de abordagens e princípios a fim de obter economicamente softwares confiáveis e que executem de forma eficiente nas máquinas reais, glossário da IEEE para terminologias em ES de 1993 [25].
- É uma disciplina da engenharia que se ocupa de todos os aspectos de produção de software, desde os estágios iniciais de especificação do sistema até a manutenção desse sistema, depois que entrou em operação. Sommerville em 2007 [26],

Em resumo, a ES está preocupada em estruturar todo o processo de produção do software, desde o primeiro contato com o cliente a manutenção exigida após o software estar em funcionamento, sempre tentando agregar técnicas, métodos, processos e outros procedimentos e mecanismos para torna o processo de desenvolvimento mais racional, científico, repetível de forma a se aproximar das engenharias tradicionais.

2.2 ESTUDOS EMPÍRICOS EM ENGENHARIA DE SOFTWARE

Pesquisas que conduzem estudos empíricos em ES deveriam ter como objetivo adquirir conhecimento sobre que tecnologias são úteis para realizar determinados tipos de tarefas e em quais ambientes [27]. Dessa forma, é possível construir um corpo de conhecimento a fim de apoiar a tomada de decisão por profissionais e pesquisadores.

O esforço empreendido para evoluir a ES e seu procedimento caminha no sentido de tornar o processo de desenvolvimento de software cada vez mais próximo do cotidiano de engenharias tradicionais, com métodos e padrões bem definidos e claros, de forma que o processo de produção do produto final seja repetível. Apesar desse esforço, atividades relacionadas com a produção de software ainda continuam dependentes da participação e criatividade humana. Desenvolvedores de software não repetem completamente a produção de um mesmo produto, dessa forma, cada solução de software desenvolvida é diferente da anterior [17], se diferenciando assim, da indústria comum que, após projetar uma vez seu produto, os produz em série.

Neste cenário onde a interferência humana é um fator importante no processo de produção, a ES, apesar de estar relacionada a ciências exatas, acaba por ter características que a torna próxima das ciências sociais e comportamentais. Tais áreas, que são influenciadas pelo comportamento humano, enfrentam os problemas relacionados aos fatores humanos com o uso de estudo empíricos para alavancar seu desenvolvimento [18].

Estudos empíricos, como a experimentação, permitem avaliar as tarefas realizadas por pessoas de forma sistemática, disciplinada, quantificável e controlada [9]. Desta forma, estudos empíricos parecem adequados para promover o desenvolvimento da ES levando em consideração fatores humanos.

Em parte da literatura, há uma aplicação dos termos estudos empíricos e experimentos de forma semelhante. Neste estudo, no entanto, um experimento é um estudo empírico, mas nem todo estudo empírico é considerado um experimento. Os experimentos, neste estudo, são representados por experimentos controlados e quasi-experimentos [9, 28, 19, 27] e são apresentados na Seção 2.3.

As informações necessárias para apoiar a decisão de qual processo, ferramenta ou tecnologia adotar, por exemplo, durante o processo de desenvolvimento de software uma empresa, são escassas. Muitas vezes os desenvolvedores de tecnologias tentam fazer com que elas sejam adotadas através de demonstrações disponibilizadas. No entanto, tais demonstrações muitas vezes ilustram nada mais que um potencial das tecnologias, de forma que é improvável que somente através de simples demonstrações sejam produzidas evidências sólidas sobre sua eficácia ou eficiência. Portanto, uma maneira de se obter as

evidências necessárias para apoiar a decisão de uma organização é colocando na prática o que se especula sobre um determinado processo, ferramenta ou tecnologia, fazendo com que as pessoas as usem durante o desenvolvimento de um software real. No entanto, as maiorias das organizações não estão dispostas a arriscar o sucesso de um projeto para avaliar uma nova tecnologia. Dessa forma, os pesquisadores e profissionais deveriam formular teorias e propor tecnologias de apoio ao desenvolvimento, e através de estudos empíricos levantarem evidências que as apoiem, construindo assim um corpo de conhecimento que sanaria o problema de falta de evidências para embasar decisões sobre o uso de processos, métodos ou tecnologias e não tomá-las baseados em intuições ou opiniões, especulações ou até mesmo de forma aleatória.

A necessidade de estudos empíricos em ES começou a ser evidenciada por Basili [3]. Nos últimos 25 anos, outros estudos têm evidenciado a importância dos estudos empíricos em ES [29, 21, 30, 9] e desenvolvido ambientes de apoio, *guidelines*, metodologias, *frameworks* e outros mecanismos que permitem projetar, planejar, executar, analisar e empacotar estudos empíricos em ES [7, 8, 9, 10, 11, 12, 13, 14].

Estudos empíricos sólidos e controlados em ES ainda são escassos [31, 32] e essa falta dificulta o surgimento e adoção de novas tecnologias, atrasando sua transferência para a indústria que segundo Redwine [33] demora em média 18 anos até ser amplamente usada. Deixando assim, nebulosa a capacidade de tomada de decisão pelos *stakeholders* que representam a indústria de software [34]. Moura [23] enumerou alguns motivos da falta desses estudos:

- A dificuldade de padronizar e encapsular todas as informações necessárias aos estudos;
- A dificuldade de controlar todas as variáveis envolvidas no desenvolvimento de software, tais como processos, tecnologias, experimentos e fatores humanos em geral;
- O envolvimento humano direto na execução dos estudos, que gera uma alta probabilidade de contaminação dos resultados, uma vez que é impossível selecionar para o estudo pessoas que tenham exatamente as mesmas características;
- Os altos custos associados à alocação de pessoas para realizar experimentos durante todo o processo de desenvolvimento de um software.

Dessa forma, fica claro a necessidade dos estudos empíricos para alavancar o amadurecimento e o desenvolvimento da área de ES.

2.3 TIPOS DE ESTUDOS EMPÍRICOS EM ENGENHARIA DE SOFTWARE

Esta seção descreve os principais tipos de estudos empíricos conduzidos em ES, que de acordo com Easterbrook [13] são:

- **Experimento Controlado:** é uma investigação de uma hipótese testável, onde uma ou mais variáveis independentes são manipuladas com a finalidade de medir seus efeitos em uma ou mais variáveis dependentes. Uma pré-condição para a realização de um experimento é uma hipótese clara, a hipótese (teoria a partir da qual o experimento é desenhado) guia todas as etapas do projeto experimental, incluindo a de decidir quais variáveis serão incluídas no estudo e como medi-las [13].
- **Estudo de Caso:** é um tipo de pesquisa empírica que oferece uma compreensão profunda de como e porque certos fenômenos ocorrem, e pode revelar os mecanismos pelos quais as relações causa-efeito ocorrem [35]. Segundo Yin [36], o estudo de caso é uma pesquisa empírica que investiga um fenômeno contemporâneo dentro de seu contexto na vida real, especialmente quando as fronteiras entre o fenômeno e o seu contexto não são claras. Estudos de caso exploratório são usados como investigações iniciais de alguns fenômenos para derivar novas hipóteses e construir teorias. Enquanto que os estudos de caso de confirmação são usados para testar teorias existentes. Uma pré-condição para conduzir um estudo de caso é uma pergunta de pesquisa clara e interessada em como e porque certos fenômenos ocorrem [13].
- **Survey:** também conhecido como uma pesquisa de campo, é usado para identificar as características de uma ampla população de indivíduos. Está estreitamente associado com o uso de questionários para coleta de dados. Entretanto, o *survey* também pode ser conduzido por meio de entrevistas estruturadas ou técnicas de registros de dados. Uma pré-condição para a realização de um *survey* é uma pergunta de pesquisa clara que pergunte sobre a natureza de uma população alvo em particular [13].
- **Etnografia:** é um tipo de estudo empírico que tem como foco o estudo de uma comunidade de pessoas através da observação. Segundo Robinson et al. [37], o objetivo da etnografia é estudar uma comunidade de pessoas para entender como os membros desta comunidade percebem suas interações sociais. Para engenharia de software, etnografia pode ajudar a compreender como as comunidades técnicas constroem uma cultura de práticas de comunicação que lhes permitem executar os trabalhos técnicos de forma colaborativa [13].

- **Pesquisa Ação:** é um tipo de estudo empírico onde os pesquisadores tentam resolver um problema do mundo real enquanto estudam simultaneamente a experiência de resolver o problema [38]. Enquanto a maioria dos tipos de pesquisa empírica tenta observar o mundo com ele atualmente existe, pesquisadores que utilizam pesquisa ação pretendem intervir nas situações estudadas com o propósito explícito de melhorar a situação [13].

2.4 ENGENHARIA DE SOFTWARE BASEADA EM EVIDÊNCIAS

Originário da virada dos anos 80 para os 90, quando as pesquisas médicas passavam por dificuldades em garantir resultados confiáveis, o paradigma baseado em evidências visou integrar as melhores evidências de pesquisas com experiências clínicas e avaliação de pacientes [12].

Embora com origem nas pesquisas médicas, o paradigma baseado em evidências é aplicável aos campos de pesquisa onde a construção do conhecimento através de revisões sistemáticas está presente, dentre elas a ES. De acordo com Kitchenham et al. [12] a Engenharia de Software Baseada em Evidências (ESBE) (do inglês *Evidence-Based Software Engineering*) busca “*prover meios pelos quais melhores evidências provenientes da pesquisa possam ser integradas com experiência prática e valores humanos no processo de tomada de decisão considerando o desenvolvimento e a manutenção de software*”.

Kitchenham et al. [12] sugere que a ESBE pode fornecer mecanismos necessários para ajudar o profissional a adotar tecnologias adequadas e evitar as inadequadas, buscando melhores práticas e procedimentos. Alguns estudos evidenciam a importância do uso do apoio da ESBA por pesquisadores para melhorar suas decisões sobre a escolha de uso de tecnologias [12, 39, 40, 41, 42].

A ESBE agrupa e avalia as evidências existentes sobre uma determinada tecnologia através de cinco etapas [43]:

1. Transformar o problema ou necessidade de informação em uma questão de pesquisa;
2. Pesquisar na literatura por melhores evidências disponíveis para responder às perguntas;
3. Avaliar criticamente as evidências, quanto a sua validade, impacto e aplicabilidade;
4. Integrar as evidências avaliadas a prática da ES;
5. Avaliar o desempenho das etapas anteriores e buscar formas de melhorá-los.

Normalmente, as etapas de 2 e 3 são realizadas através de uma Revisão Sistemática da Literatura (RSL) (do inglês *Systematic Literature Review*), que é um dos principais métodos empregado pela ESBE. Elas são classificadas como estudos secundários, já que, dependem dos estudos primários conduzidos para poder agregar evidências e construir conhecimento [39, 40, 41, 42].

Segundo Petersen et al. [44], com o amadurecimento de uma área de pesquisa, o número de estudos e resultados crescem significativamente. É importante então, que a comunidade mantenha um rastreamento dos avanços na forma de RSLs Recentemente tal estudo está sendo amplamente usado [41, 45, 46]. A literatura diferencia alguns tipos de RSLs [47]:

- **RSLs concencionais** [47]: agregam resultados sobre a eficácia de um determinado tratamento, intervenção ou tecnologia, e estão relacionadas a questões de pesquisa como: *O tratamento X sobre uma população Y é mais eficiente em obter o resultado W no contexto K em comparação com o tratamento Z?*
- **Estudos de Mapeamento Sistemático** [48]: objetiva identificar todas as pesquisas relacionada a um tópico específico, por exemplo, responder questões que são mais amplas (exploratórias) sobre tendências de pesquisa. Uma questão típica é: *O que nós sabemos sobre o Tema X?*

De acordo com Kitcheham et al. [41], as principais diferenças entre um EMS e uma RSL convencional são:

- Geralmente, EMS são guiados por questões amplas, mais exploratórias, e por muitas vezes elabora múltiplas questões de pesquisa;
- Os termos de pesquisa para um EMS são menos focados do que os usados em uma RSL, com isso são mais suscetíveis a encontrar um número maior de estudos nas buscas. Para um EMS, no entanto, isto é um problema menor do que encontrar um grande número de estudos durante a etapa de busca em uma RSL, já que o objetivo de um EMS é ter uma ampla cobertura e não restringir o foco;
- O processo de extração de dados para um EMS é também muito mais amplo do que do processo de uma RSL e pode ser chamado mais especificamente de classificação ou estágio de categorização. A proposta deste estágio é classificar artigos com detalhes suficientes para responder as amplas (exploratórias) questões de pesquisa e identificar artigos para revisões anteriores, sem ser uma tarefa demorada;

- A etapa de análise de um EMS trata da síntese dos dados para responder as questões de pesquisa propostas. É pouco provável que incluam técnicas de análise profundas, com a meta-análise ou síntese narrativa, mas resultados das contagens e resumos. As representações gráficas das distribuições de estudos por tipo de classificação pode ser um mecanismo de comunicação eficaz;
- A divulgação dos resultados de um EMS pode ser mais limitada do que uma RSL; limitada aos que requisitam o trabalho e a publicações acadêmicas, com o objetivo de influenciar o rumo das pesquisas primárias.

RSLs e EMSs surgiram como duas importantes ferramentas para agregar e construir conhecimento em ES, elas apresentam as seguintes vantagens [41]:

- A metodologia bem definida torna menos provável que os resultados encontrados na literatura sejam tendenciosos, embora não proteja contra o viés dos estudos primários selecionados;
- Elas podem fornecer informações sobre os efeitos de algum fenômeno através de uma ampla gama de configurações e métodos empíricos. Se os estudos fornecem resultados consistentes, a revisão sistemática prover evidências que o fenômeno é robusto e transferível. Se os resultados fornecidos são inconsistentes, as fontes de variação podem ser estudadas;
- No caso de estudos quantitativos, é possível combinar dados usando técnicas de meta-análise. Isso aumenta a probabilidade de detectar efeitos reais que pequenos estudos individuais não são capazes de detectar.

Nesta dissertação será realizado um EMS coma finalidade de mapear mecanismos que guiem estudos empíricos em ES. Alguns motivos para justificar a realização de um EMS são apresentados por Askey e O'Malley [48]:

- Para examinar a extensão, alcance e natureza dos fenômenos de investigação. Este tipo de revisão é útil para mapeamento de áreas de estudo, onde é difícil visualizar a gama de materiais que possam estar disponíveis;
- Para determinar a necessidade de realizar uma revisão sistemática convencional. Nestes casos, um mapeamento preliminar da literatura pode ser realizado para identificar se é viável ou relevante executar uma revisão sistemática convencional;

- Para resumir e divulgar resultados de pesquisas. Esse tipo de estudo de escopo pode descrever em detalhes os resultados e a abrangência da pesquisa em determinadas áreas de estudo, proporcionando assim um mecanismo de síntese e divulgação dos resultados da investigação;
- Para identificar as lacunas de pesquisa na literatura existente. Este tipo de estudo de escopo elabora conclusões, a partir da literatura existente, sobre o estado global da área de pesquisa investigada. Especificamente concebido para identificar lacunas na base de evidências onde não há pesquisas sendo conduzidas, o estudo de escopo pode também resumir e disseminar resultados de pesquisas, bem como, identificar a relevância de uma revisão sistemática convencional em uma área de estudo específica. No entanto é importante notar que identificar lacunas na literatura através do estudo de escopo não é necessariamente identificar lacunas na pesquisa em si, ou seja, identificar se uma pesquisa tem uma qualidade pobre, uma vez que a avaliação da qualidade não faz parte da missão de um estudo de escopo.

Uma tarefa que é considerada fundamental na execução de um EMS é a elaboração de um protocolo detalhado de que descreva o processo e os métodos que serão aplicados na pesquisa sistemática [39]. O protocolo, dessa forma, se torna um item essencial para diminuir a possibilidade do pesquisador enviesar a pesquisa e necessita ser pré-definido. Além de que, torna possível a replicação do estudo por outros pesquisadores. O protocolo é discutido de forma breve no Capítulo 3 e apresentado de forma completa no Apêndice C.

Segundo Travassos [42], a etapa de planejamento de uma revisão sistemática, deve contemplar os seguintes passos:

- Listar os objetivos da pesquisa;
- Formular as questões de pesquisa (strings de busca elaboradas);
- Apresentar os métodos que serão utilizados para a execução e análise dos dados obtidos;
- As fontes e seleção de estudos devem ser planejadas;
- Um protocolo deve ser definido.

Uma vez produzido o protocolo a revisão sistemática pode ser iniciada, no caso um EMS. De acordo com Travassos [42], os passos a serem seguidos durante a execução da revisão sistemática são:

- Realização das buscas nas fontes definidas. O processo deve ser transparente, passível de repetição e toda sua execução documentada;
- Seleção dos estudos primários de acordo com os critérios de inclusão e exclusão definidos;
- Extração dos dados, desde informações gerais sobre os estudos encontrados e os selecionados até as que são julgadas importantes para responder as questões de pesquisa.

Além dos passos citados anteriormente, também é interessante realizar uma busca manual nos principais anais de conferências e periódicos relevantes para o tema estudado. A busca manual pode ser importante para dar mais confiança ao pesquisador que conduz a revisão sistemática, ampliar a cobertura da revisão, captar estudos que ainda não foram indexados pelos engenhos de busca das bibliotecas digitais, devido a publicação recente, e identificar estudos relevantes que não usam os termos da *string* usada na busca automatizada.

O EMS finda por apresentar os mapas das evidências que categorizam a área de pesquisa estudada e a avaliação dos mesmos. Os resultados são apresentados em um nível de granularidade mais elevado, de forma que respondam as questões de pesquisa, permita a identificação de agrupamentos e "desertos" de evidências para direcionar o foco de futuras RSLs e identificar áreas onde mais estudos primários são necessários [41].

2.5 RESUMO

Neste capítulo foi apresentado o embasamento teórico utilizado durante esta pesquisa. Foram apresentados conceitos como os de ES e o uso de pesquisa empírica neste campo. Alguns tipos de estudos empíricos, considerados importantes para a ES foram introduzidos. Por fim, houve a apresentação de uma área de pesquisa em ES denominada ESBE, bem como das principais métodos de pesquisa usados nela.

CAPÍTULO 3

METODOLOGIA

Seguir uma metodologia científica é indispensável em um trabalho acadêmico, pois ela traz confiança ao estudo e torna possível sua replicação, de forma independente, por outros pesquisadores. Este capítulo apresenta a abordagem metodológica usada nesta pesquisa e está organizado em duas partes:

- **Classificação da Pesquisa:** apresentação da classificação da pesquisa juntamente com o quadro metodológico;
- **Ciclo da Pesquisa:** descrição das etapas da pesquisa, bem como do processo de mapeamento sistemático.

3.1 CLASSIFICAÇÃO DA PESQUISA

Esta pesquisa utiliza o método de abordagem **indutivo** baseado em dados de natureza **qualitativa**. Os método de procedimento utilizado foi o **estudo de mapeamento sistemático**, que é um tipo de revisão sistemática [41] .

A Tabela 3.1 apresenta o quadro metodológico do trabalho.

Tabela 3.1 Classificação da Pesquisa.

Quadro Metodológico	
Método de Abordagem	Indutivo
Método de Procedimento	Estudo de Mapeamento Sistemático
Natureza dos dados	Qualitativa

De acordo com Marconi e Lakatos [49] o método de abordagem indutivo é:

“um processo mental por intermédio do qual, partindo de dados particulares, suficientemente constatados, infere-se uma verdade geral ou universal, não contida nas partes examinadas.”

Para toda indução devem ser considerados três elementos essenciais:

- Observação dos fenômenos, com a finalidade de descobrir as causas de sua manifestação;
- Descoberta da relação, por intermédio da comparação, com a finalidade de descobrir a relação constante existente entre eles;
- Generalização da relação entre os fenômenos e fatos semelhantes.

Os métodos de procedimento, que é a etapa mais concreta no processo de investigação, definido para esta pesquisa é o Estudo de Mapeamento Sistemático (EMS), do inglês *Systematic Mapping Study*, que é uma forma de avaliar e interpretar todas as pesquisas disponíveis referentes a uma questão de pesquisa em particular, área temática, ou fenômeno de interesse. A escolha de um mapeamento sistemático, ao invés de uma revisão sistemática tradicional, para esta pesquisa foi baseada na natureza exploratória das questões de pesquisa.

Por fim, no que diz respeito ao uso de pesquisa qualitativa, Marconi e Lakatos [49] destacam que a escolha por esse tipo de método de pesquisa possuem vantagens quando comparado aos demais, pois o mesmo preocupa-se em analisar e interpretar aspectos mais profundos, descrevendo a complexidade do comportamento humano, fornecendo dessa forma, análises mais detalhadas sobre as investigações, hábitos, atitudes, tendências de comportamento, etc. Além disso, os métodos qualitativos são capazes de prover informações mais exploratórias e ajudam a refinar as proposições para que melhor se ajustem aos dados.

No entanto, o uso de pesquisa qualitativa também traz algumas desvantagens. Dentre elas pode-se destacar o maior esforço que o pesquisador tem que despender em relação ao método quantitativo e a maior dificuldade para resumir os achados. Além disso, resultados qualitativos frequentemente são considerados mais nebulosos do que os quantitativos, especialmente em comunidades mais técnicas como na engenharia de software.

3.1.1 Classificação segundo Cooper

A fim de complementar a descrição da metodologia apresentada, será usado a taxonomia proposta por Cooper [50] para classificar revisões sistemáticas. A taxonomia proposta usa cinco características para fazer a classificação: foco, objetivo, perspectiva, cobertura, organização e audiência. A Tabela 3.2, apresenta esta classificação.

Cada característica da classificação de Cooper [50] será apresentada a seguir.

- *Foco*: está relacionado ao interesse principal do pesquisador, e por sua vez tem quatro categorias: Resultados de Pesquisa, Métodos de Pesquisa, Teorias e Práticas

Tabela 3.2 Classificação segundo Cooper.

Características	Categorias
Foco	Resultados de Pesquisa Métodos de Pesquisa Práticas ou Aplicações
Objetivo	Integração
Perspectiva	Representação Neutra
Cobertura	Exaustiva
Organização	Conceitual
Audiência	Pesquisadores Especializados Profissionais

ou Aplicações. Esses quatro focos não excludentes entre si, e geralmente uma revisão tem mais de um foco, cada qual variando no seu nível de atenção;

- *Objetivo*: diz respeito aos resultados que o pesquisador almeja alcançar. Podem ser classificados como: Integração, Crítica e Identificação de Problemas Centrais. Como a característica citada no item anterior, essas categorias também não são excludentes entre si. O objetivo mais comum encontrado em uma revisão é o da integração, pois corresponde a síntese da literatura existente em uma área. Também pode ser uma análise crítica da literatura, quando a intenção do pesquisador é demonstrar que os resultados de estudos anteriores são injustificáveis e por fim há a identificação de problemas centrais, em que a intenção é fornecer sugestões sobre como os problemas e controvérsias em uma área se tornaram evidentes;
- *Perspectiva*: são as posições assumidas pelo pesquisador nas discussões da pesquisa. São duas categorias: Representação Neutra ou Exposição de Posicionamento. No primeiro caso, a intenção do pesquisador é apresentar argumentos e evidências sobre as diferentes teorias, métodos, problemas, ou resultados, de forma a revelar os trabalhos produzidos na literatura. Com relação à segunda perspectiva, o pesquisador tem uma postura mais ativa com a intenção de apresentar e defender um ponto de vista;
- *Cobertura*: descreve a decisão que o pesquisador tomou sobre a forma de realizar as buscas e a inclusão dos trabalhos relevantes para sua pesquisa, ou seja, como o revisor pesquisa a literatura e como ele toma decisão sobre a qualidade do material. São quatro os tipos de cobertura: Exaustiva, Exaustiva com Seleção de Citação, Representativa e Central ou Essencial. Na cobertura exaustiva, a intenção do pes-

quisador é coletar toda ou quase toda a literatura disponível na área. No segundo tipo de cobertura a exaustão também é a meta, porém somente uma amostra selecionada é apresentada no trabalho. Já na cobertura representativa o pesquisador escolhe a amostra que representa a área de pesquisa total. E por fim, na cobertura central ou essencial, o pesquisador seleciona os trabalhos que focam em esforços importantes e iniciais que direcionam um campo de pesquisa;

- *Organização*: diz respeito a organização do trabalho e dos resultados. Podem ser apresentados de três formas: Historicamente, onde os temas são apresentados em ordem cronológica do aparecimento na literatura, Conceitualmente, que apresenta trabalhos relacionados onde as mesmas idéias aparecem juntas, ou Metodologicamente, onde os trabalhos que empreguem métodos semelhantes são agrupados em subtemas. Pesquisadores que estejam fazendo uma revisão podem combinar as organizações abordando, por exemplo, trabalhos historicamente dentro de um quadro conceptual e metodológico;
- *Audiência*: apresenta a audiência da pesquisa, ou seja, o público-alvo. Revisores podem escrever para Pesquisadores Especializados, Pesquisadores em Geral, Profissionais, ou para o Público em geral. A distinção do público-alvo é que vai determinar o estilo de escrita e emprego de jargões e conceitos específicos da área.

3.2 CICLO DA PESQUISA

As etapas que compõem o ciclo de vida utilizado nesta pesquisa é ilustrado pela Figura 3.1.

Após identificar a necessidade de mecanismos para guiar estudos empíricos no trabalho de Moura [23], foi realizada uma pesquisa bibliográfica tradicional sobre o uso de pesquisa empírica em engenharia de software com a finalidade de obter embasamento teórico sobre o assunto.

Através de observação e análise dos estudos encontrados ficou evidente a importância do tema para o desenvolvimento da engenharia de software. Além disso, não foram encontrados estudos sistemáticos que consolidem e sumerizem o conhecimento da área. Portanto, ficou clara a necessidade de executar um mapeamento sistemático da área. A partir deste ponto, o foco desta pesquisa foi definido e expresso em questões de pesquisa que serão apresentadas na Seção 3.2.2.

Posteriormente, um protocolo para execução de um EMS foi desenvolvido e executado para coletar evidências com o intuito de responder as perguntas de pesquisa propostas.

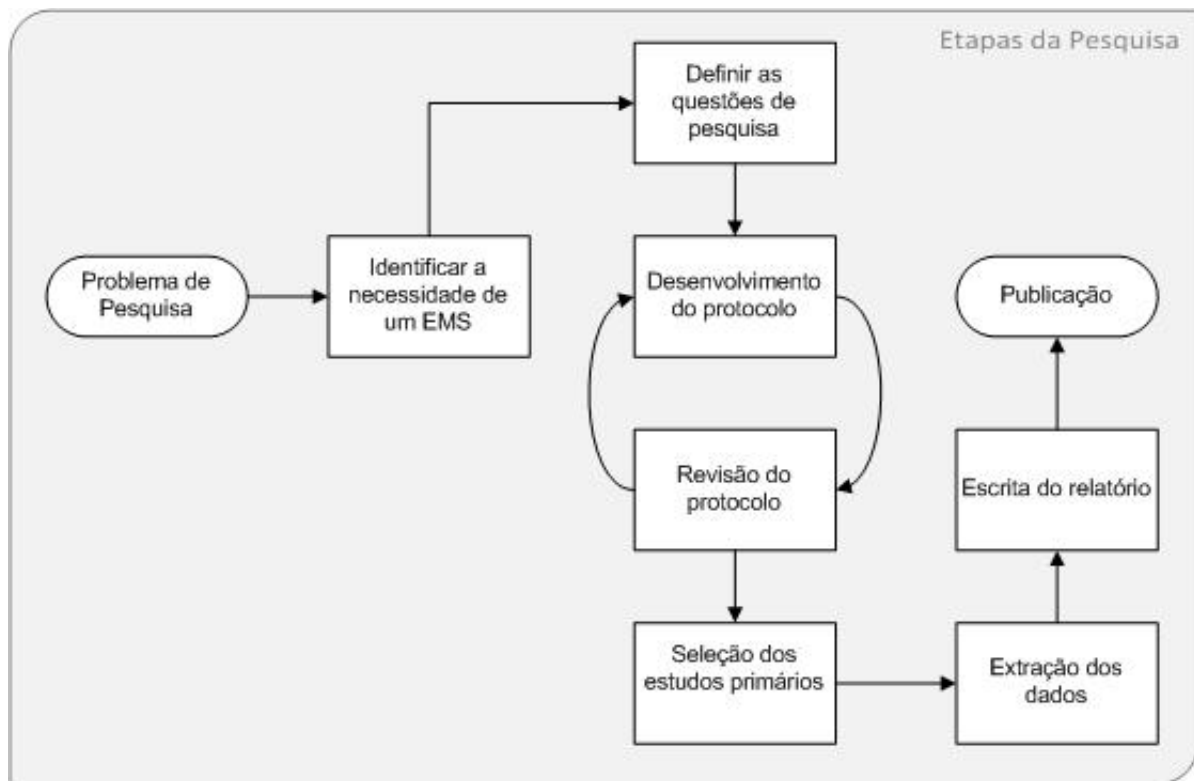


Figura 3.1 Etapas da Pesquisa

Por fim, o relatório do mapeamento escrito, e desta forma, visa prover uma contribuição à comunidade científica com a agregação do conhecimento existente em engenharia de software de como guiar estudos empíricos. Além disso, este estudo analisa e classifica os estudos disponíveis e aponta lacunas existentes.

3.2.1 Mapeamento Sistemático

Esta seção apresenta, de forma breve, alguns pontos do protocolo seguido pelo mapeamento sistemático conduzido por este estudo no campo da engenharia de software empírica. O protocolo integral encontra-se disponível no Apêndice C.

Esta pesquisa conduziu o mapeamento com a intenção de encontrar e analisar o maior número de trabalhos primários relevantes e reconhecidos na área de ES a fim responder as perguntas de pesquisa, definidas a seguir na Seção 3.2.2.

De acordo com Kitchenham et al. [41], as perguntas de pesquisa de um estudo de mapeamento sistemático são muito mais amplas e gerais (normalmente de caráter exploratório), em contraste as que são elaboradas em revisões sistemáticas da literatura, que

devem ser bem mais precisas. Para guiar a construção deste protocolo de pesquisa foram usadas as instruções e conselhos apresentadas no *guideline* definido por Kitchenham et al. [41].

Segundo Kitchenham et al. [12] revisões sistemáticas e mapeamentos sistemático de estudos tem em comum:

1. Identificação da necessidade de executar a revisão sistemática;
2. Elaboração da questão de pesquisa;
3. Uma abrangente e exaustiva busca por estudos primários;
4. Avaliação de qualidade dos estudos incluídos;
5. Identificação dos dados necessários para responder a pergunta de pesquisa;
6. Extração dos dados;
7. Resumo e síntese dos resultados dos estudos (meta-análise);
8. Interpretação dos resultados para determinar sua aplicabilidade;
9. Escrita do relatório.

3.2.2 Escopo e Questões de Pesquisa

Com o objetivo de delinear o escopo da pesquisa e de identificar os elementos que vieram a fazer parte das questões de pesquisa, foi utilizada uma estrutura citada por Kitchenham et al. [41], que baseada no trabalho de Petticrew et al. [47], recomenda considerar as questões de pesquisa a partir da seguinte estrutura denominada PICOC:

- **População** (*Population*): Pesquisas que conduzem estudos empíricos em engenharia de software.
- **Intervenção** (*Intervention*): uso de mecanismos para guiar estudos empíricos.
- **Comparação** (*Comparison*): não se aplica.
- **Resultado** (*Outcomes*): Um documento que mapeia os mecanismos para guiar estudos empíricos em engenharia de software.
- **Contexto** (*Context*): indústria e academia.

O item Comparação (*Comparison*) não foi utilizado, uma vez que o estudo não realiza comparações entre os mecanismos para guiar estudos empíricos. O item da estrutura denominado contexto (*Context*) é normalmente utilizado para definir o contexto que ocorre a comparação. No entanto, aqui será utilizado para definir o contexto admitido para os mecanismos encontrados.

Após a definição da estrutura PICOC as perguntas que guiaram o estudo foram definidas.

- **Questão 1:** Quais são os mecanismos existentes para guiar estudos empíricos em engenharia de software?
- **Questão 2:** Os mecanismos encontrados são utilizados para guiar que tipo de estudo empírico?
- **Questão 3:** Quais elementos, fases, etapas compõem os mecanismos utilizados para guiar estudos empíricos?

3.2.3 Estratégia de Busca

A construção da *string* de busca utilizada nas bibliotecas digitais selecionadas seguiu uma estratégia baseada em [51], que consiste nos seguintes passos:

- Derivar a partir das questões de pesquisa as principais palavras-chaves a partir da identificação da população, intervenção, comparação (quando for o caso), resultados e contexto;
- Procurar por palavras chaves em artigos relevantes já consultados em uma revisão informal;
- Identificar sinônimos e termos alternativos as palavras chaves;
- Consultar especialistas na área;
- Usar o conector booleano OR para incorporar palavras alternativas e sinônimas.
- usar o conector booleano AND para ligar palavras chaves
- Verificar a string de busca construída realizando buscas piloto e comparando os resultados obtidos com uma lista de estudos primários já conhecidos;

Como resultado da estratégia supracitada foi obtido à seguinte string de busca:

(framework OR method OR methodology OR guideline OR process OR report OR reporting OR support OR supporting OR checklist OR tutorial) AND (empirical OR experiment OR experimentation OR experimental OR survey OR “case study” OR “action research” OR ethnography OR “controlled experiment” OR quasi-experiment OR “quasi experiment”) AND (“software engineering”)

3.2.4 Processo de Busca

O processo que foi usado para procurar por estudos primários incluiu buscas automatizadas através de engenhos de busca das bibliotecas digitais, onde foi utilizada a *string* de busca formulada, e também com a execução de buscas manuais em anais de eventos e periódicos relevantes, com o objetivo de ampliar a cobertura da pesquisa e dar mais segurança ao pesquisador. No caso da busca manual, o que determinou o período de tempo que limitou a busca foi a disponibilidade do material na internet.

As bibliotecas digitais utilizadas na busca automatizada foram:

- IEEE Computer Society Digital Library;
- ACM;
- Science Direct;
- Scopus;
- EI Compendex.

As conferências utilizadas na busca manual foram:

- Empirical Software Engineering and Measurement (ESEM), no período de 2007 a 2010;
- Evaluation and Assessment in Software Engineering (EASE), no período de 2006 a 2010.

E por fim, o periódico utilizado na busca manual foi:

- Empirical Software Engineering Journal (ESEJ), no período de 1997 a 2010.

3.2.5 Critérios de Inclusão/Exclusão e Procedimentos de Seleção

O procedimento adotado para seleção dos estudos primários desta pesquisa foi dividido em duas etapas. Na primeira etapa, dois pesquisadores analisaram o conjunto de estudos, obtidos através das bibliotecas digitais, lendo o título, resumo e palavras-chave. Ao fim desta etapa um conjunto de potenciais estudos primários foi gerado. Na segunda etapa de seleção foram obtidos todos os artigos completos do conjunto de estudos potencialmente relevantes, resultante da primeira etapa, e este conjunto foi dividido em duas metades, de modo que cada metade foi avaliada por dois pesquisadores a fim de gerar o conjunto final de estudos primários do mapeamento. O procedimento seguido foi sempre supervisionado pelo orientador da dissertação através de reuniões de resolução de conflito ao fim de cada etapa.

A inclusão de um trabalho no mapeamento se dá pela sua relevância em relação às questões de investigação. Os critérios de inclusão adotados pelo estudo foram:

- Estudos que definem um mecanismo para guiar estudos empíricos;
- Estudos que estão relacionados a engenharia de software.

Os critérios de exclusão adotados pelo estudo foram:

- Estudos irrelevantes para a pesquisa, ou seja, aqueles que não definem nenhum mecanismo para guiar um estudo empírico;
- Estudos que são repetidos, se um estudo estiver disponível em diferentes fontes, a primeira pesquisa será a considerada;
- Estudos Duplicados, no caso de haver dois ou mais estudos semelhantes apenas o mais completo será considerado, exceto em situações onde há informações adicionais.
- Estudos incompletos como resumos ou resumos expandidos e apresentações (slides).

3.2.6 Extração dos dados

Os dados usados como evidências para responder as perguntas de pesquisa deste estudo foram coletados através do formulário de coleta disponível no Apêndice C.

3.2.7 Estratégia de Síntese dos Dados

A contribuição idealizada por este estudo é gerada a partir da síntese dos dados obtidos através do mapeamento sistemático realizado. O procedimento para a síntese dos dados se inicia com a marcação das passagens retiradas dos estudos selecionados (dados qualitativos), a partir dos formulários apresentados no Apêndice C, que fornecem algum tipo de informação relevante para responder as perguntas de pesquisa. Cada um dos trechos retirados dos estudos primário é associado a um código que identifica o tipo ou categoria de informação que a passagem prover. Os códigos devem seguir um padrão para posteriormente serem agrupados. Por exemplo, nesta pesquisa foram extraídos passagens que indicam o tipo de mecanismo desenvolvido, qual tipo de estudo o mecanismo guia e a forma que está estruturado o guia. Os códigos deste estudo seguem a seguinte formatação:

- **ID_TM**;
- **ID_TEG**.

Onde:

- **ID** = **ID**entificador do estudo primário;
- **TM** = **T**ipo do **M**ecanismo encontrado para guiar os estudos empíricos em engenharia de software;
- **TEG** = **T**ipo do **E**studo **G**uiado pelo mecanismo desenvolvido.

Assim, pode-se agrupar os trechos das pesquisas analisadas em grupos de acordo com os códigos utilizados. O primeiro padrão de código identificou os trechos dos estudos que indicava o tipo de mecanismo desenvolvido, o segundo código identificou os trechos que apontavam o tipo de estudo guiado pelo mecanismo e por fim o terceiro código foi utilizado para identificar os trechos dos estudos que descreviam a estrutura do mecanismo desenvolvido. A aplicação deste procedimento pode ser vista no Capítulo 4.

Desta forma, seguindo o procedimento descrito e utilizando do auxílio de planilhas eletrônicas, foram mapeadas as evidências dos estudos selecionados e sintetizados os resultados apresentados no próximo capítulo. Todo processo foi realizado de forma que as informações possam ser rastreadas, possibilitando identificar os trabalhos que proveram os dados.

3.3 RESUMO

Neste capítulo foi descrito a metodologia utilizada nesta pesquisa, como foi estruturada, conduzida e as razões de uso dos procedimentos ou métodos. Além de uma breve descrição do protocolo usado para guiar a execução do EMS proposto pelo estudo. Portanto, dessa forma espera-se atingir o rigor necessário para obter validade científica, essencial para a confiabilidade dos resultados deste estudo.

CAPÍTULO 4

RESULTADOS

Este capítulo apresenta os resultados do EMS obtidos através da execução do protocolo definido no capítulo anterior. Os resultados estão estruturados em três componentes, descritos a seguir.

- **Extração e análise dos dados:** apresentação dos dados gerais sobre o estudo, tais como: quantidade de estudos retornados na busca automatizada e manual, processo de seleção com o número final de estudos primários, a distribuição das datas de publicação de cada um ao longo dos anos, dentre outros;
- **Mapeamento das evidências:** apresentação de trechos transcritos dos estudos primários selecionados, coletados como evidências, ou seja, os resultados para as questões de pesquisa Q1, Q2 e Q3;
- **Discussão sobre os resultados:** apresentação da análise dos principais resultados obtidos com o estudo.

4.1 EXTRAÇÃO E ANÁLISE DOS DADOS

O mapeamento sistemático de estudos foi realizado de acordo com o protocolo apresentado resumidamente no Capítulo 3 e por completo no Apêndice A. A *string* de busca definida juntamente com as cinco bibliotecas digitais eleitas para serem usadas no estudo, retornaram 6439 trabalhos, sendo, 729 resultado da busca na *IEEE*, 374 na *ACM*, 307 na *Science Direct*, 261 na *Scopus* e 4768 na *EI Compendex*. O gráfico na Figura 4.1 ilustra a participação das bibliotecas digitais no montante de estudos encontrados.

Além da busca realizada nas bibliotecas digitais, buscas manuais também foram empreendidas com o objetivo de aumentar o alcance do estudo. As buscas manuais foram realizadas em conferências e periódicos de relevância para a área estudada, foram eles: a conferência *Empirical Software Engineering and Measurement (ESEM)*, no período de 2007 a 2010, a conferência *Evaluation and Assessment in Software Engineering (EASE)*, no período de 2006 a 2010 e o periódico *Empirical Software Engineering Journal (ESEJ)*,

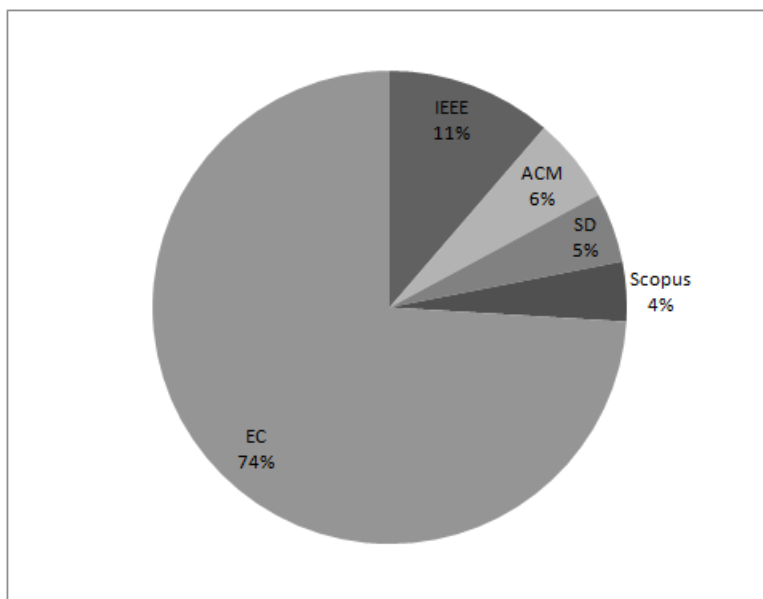


Figura 4.1 Resultado da busca automatizada.

no período de 1997 a 2010. O montante de estudos que cada conferência ou periódico tem a oferecer é conhecido através de uma simples contagem dos estudos em seus anais (se for conferência) ou nas edições (se for periódico), o resultado da contagem foi 662, dos quais, 236 estudos são do *ESEM*, 70 do *EASE* e 356 do periódico *ESEJ*. O gráfico na Figura 4.2 ilustra a participação das conferências e do periódico no montante de estudos encontrados na busca manual.

O resultado final dos dois tipos de busca foi um espaço de busca de estudos primários cujo montante é de 7101 estudos, onde aproximadamente 9% é resultado da busca manual e 91% da busca automatizada. O gráfico na Figura 4.3 ilustra esse resultado.

Como pode ser observado, o número de estudos retornados pela busca automatizada e manual foi alto. Este problema é compartilhado pelos estudos secundários sistemáticos (revisões sistemáticas da literatura e mapeamentos sistemáticos de estudos) que utilizam processos automatizados de busca em função das características e funcionalidades dos engenhos de busca [41].

Após a leitura do título e resumo, cada um por dois revisores, o alto número de estudos retornados inicialmente foi drasticamente reduzido de 7101 para 114 (uma redução de aproximadamente 98%). Dessa forma foi viabilizado uma melhor análise do material descartando os estudos identificados como irrelevantes de acordo com os critérios de inclusão e exclusão apresentado no Capítulo 3. O gráfico na Figura 4.4 mostra como os

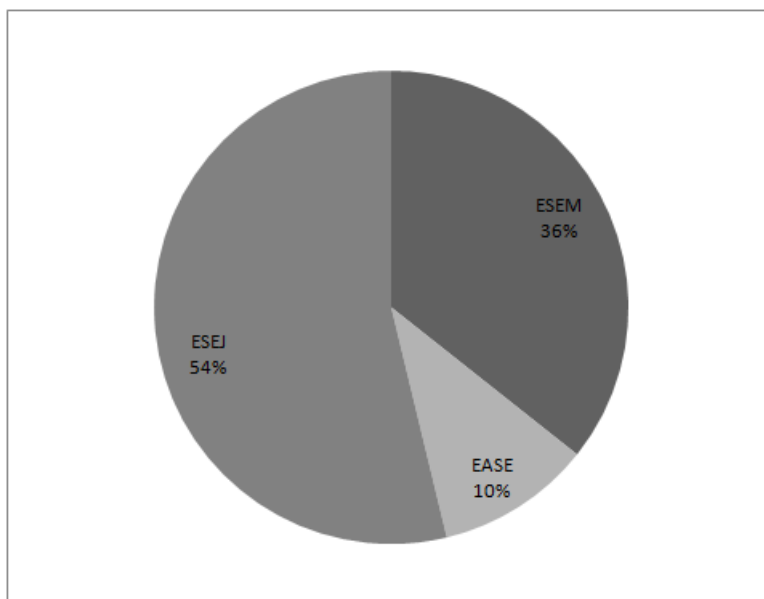


Figura 4.2 Resultado da busca manual.

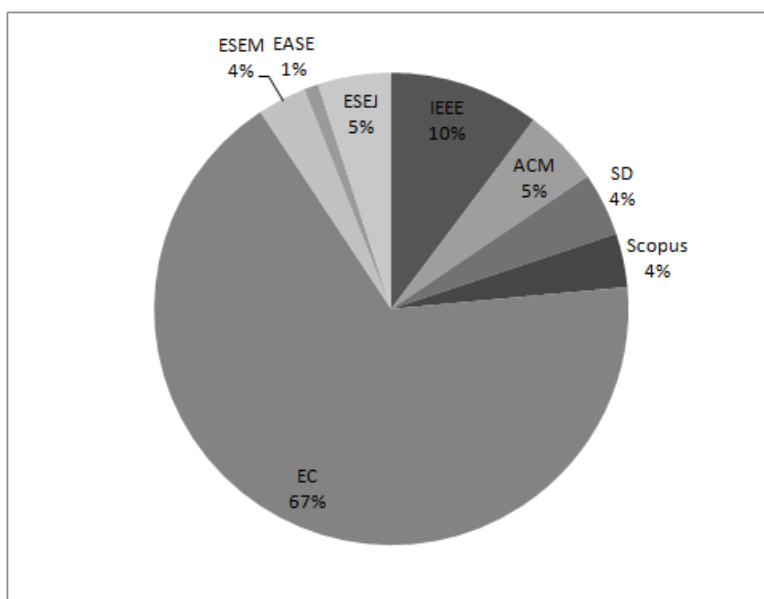


Figura 4.3 Resultado das buscas automatizada e manual juntas.

trabalhos selecionados como potencialmente relevantes estão distribuídos.

Com os potenciais estudos selecionados, todos os artigos foram lidos completamente, por duas duplas (composta por três revisores), com o objetivo de selecionar os estudos primários que compõem este mapeamento. Usando os critérios de inclusão e exclusão,

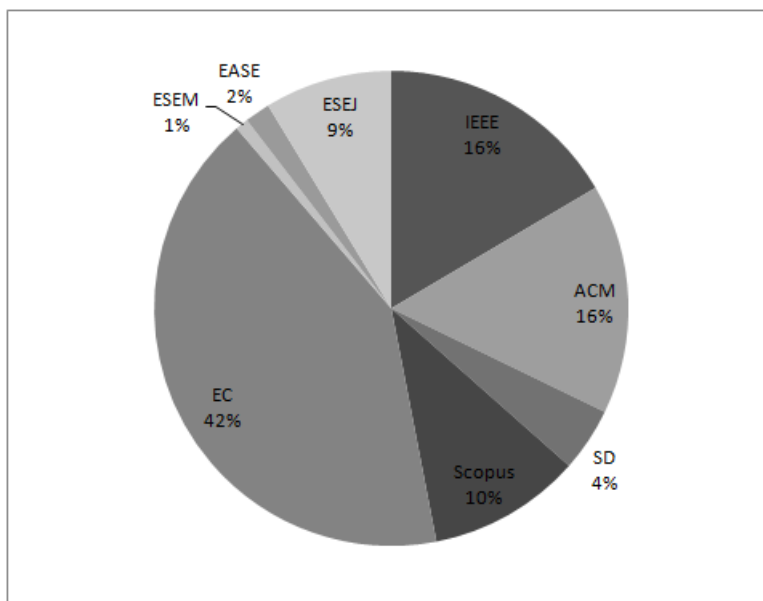


Figura 4.4 Resultado seleção de trabalhos potencialmente relevantes para o estudo.

chegou-se a 23 estudos primários (disponíveis no Apêndice A. Dessa forma os 92 estudos considerados potencialmente relevantes na etapa anterior foram excluídos por serem repetidos, eram irrelevantes, não eram artigos completos ou simplesmente não se teve acesso ao artigo. O gráfico na Figura 4.5 mostra a distribuição dos estudos primários selecionados.

Uma observação interessante pode ser feita em relação a quantidade de estudos retornados pelo *EI Compendex*, dominante na busca inicial (levando em consideração a busca manual) com 67% dos trabalhos retornados e 42% após a leitura do título e abstract dos trabalhos. No entanto, o mesmo ficou com uma fatia de apenas 22% dos trabalhos primários selecionados, sendo superado pelo *IEEE* com 30% e *ACM* com 26%. Isso pode ser explicado pelo fato do *EI Compendex* ter sido o último engenho de busca a ser utilizado e por isso teve uma quantidade relevante de trabalhos considerados repetidos ou duplicados, embora que nem todos os trabalhos encontrados pelo engenho de busca da *ACM* e do *IEEE* tenham sido encontrados pelo *EI Compendex*. O fato de iniciar a busca pelo *IEEE* e *ACM* foi proposital, pois foi dada preferência a começar por esses dois mecanismos originais de publicação de trabalhos. Contudo, isso é interessante perceber mecanismos como *Ei Compendex* são cada vez mais poderosos e no futuro tendem concentrar todas as consultas em apenas um local, reduzindo o esforço do pesquisador nesta etapa. A Tabela 4.1 mostra de forma sintética a evolução do processo de seleção

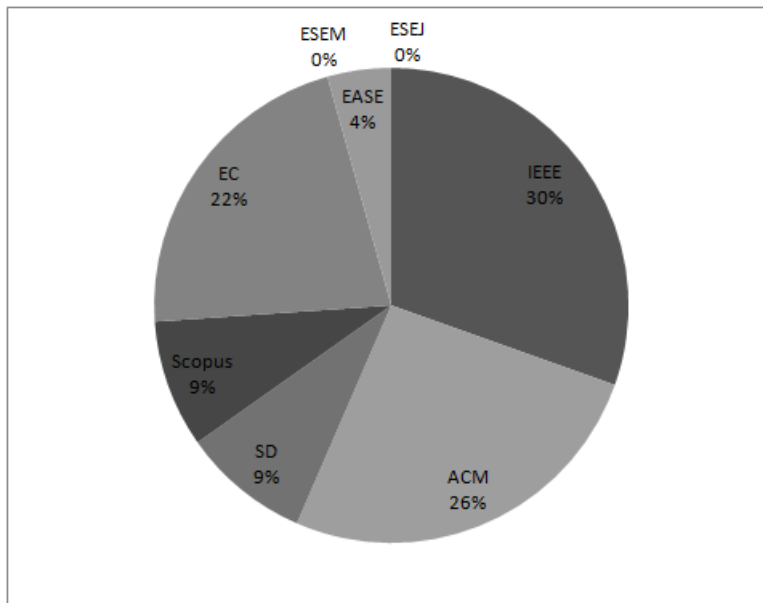


Figura 4.5 Resultado seleção de trabalhos primários que compõem o estudo.

dos estudos primários.

Tabela 4.1 Evolução do processo de seleção.

Seleção de Estudos Primários							
Fontes	Estudos Retornados	#1 Seleção Estudos Potencialmente Relevantes	#2 Seleção				Estudos Primários
			Excluídos				
			Irrelevantes	Repetidos/Duplicados	Incompleto	Sem Acesso	
IEEE	729	19	9	0	3	0	7
ACM	374	17	4	3	3	0	7
SD	307	5	2	0	0	1	2
Scopus	261	12	3	4	3	0	2
EC	4768	48	19	19	4	2	4
ESEM	236	1	0	1	0	0	0
EASE	70	2	1	0	0	0	1
ESEJ	356	10	2	6	2	0	0
TOTAL	7101	114	40	33	15	3	23

Apesar da busca automatizada em engenhos de busca não ter sido limitada a um período específico, todos os estudos encontrados estão entre 1986 e 2009, sendo a maioria concentrados na última década. Isso evidencia que pesquisas visando viabilizar a melhor execução de estudos empíricos específicos para engenharia de software têm aumentado na última década. O gráfico da Figura 4.6 ilustra a concentração dos estudos por ano.

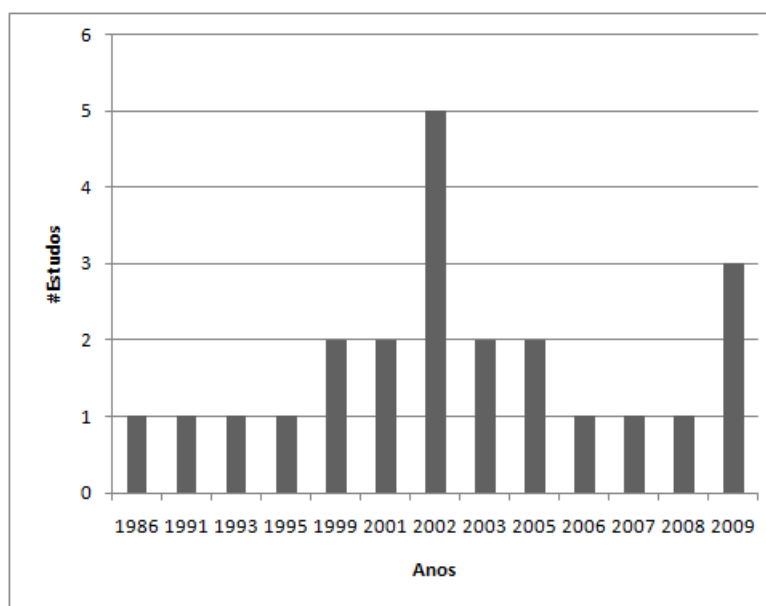


Figura 4.6 Distribuição dos estudos ao longo dos anos.

O mapeamento sistemático contabilizou 55 autores nos 23 estudos primários selecionados, dentre os quais estão presente em mais de um artigo: Victor R. Basili, Richard W. Selby, Barbara A. Kitchenham, Shari Lawrence Pfleeger e Lesley M. Pickard. Os pesquisadores tem origem em 11 países diferentes, conforme detalhamento da Figura 4.7. A soma das publicações de cada país supera a quantidade de estudos selecionados porque alguns estudos foram desenvolvidos em cooperação entre dois ou mais pesquisadores que tem origem em instituições de países distintos.

Esta seção apresentou dados gerais coletados no mapeamento sistemático. Tais informações podem ser usadas como referência para estudos semelhantes ou replicações para confirmar ou refutar os resultados apresentados.

4.2 MAPEAMENTO DAS EVIDÊNCIAS

Nessa seção, as evidências capturadas através dos estudos primários selecionados são apresentadas respondendo as perguntas de pesquisa que guiaram o estudo. A seção é

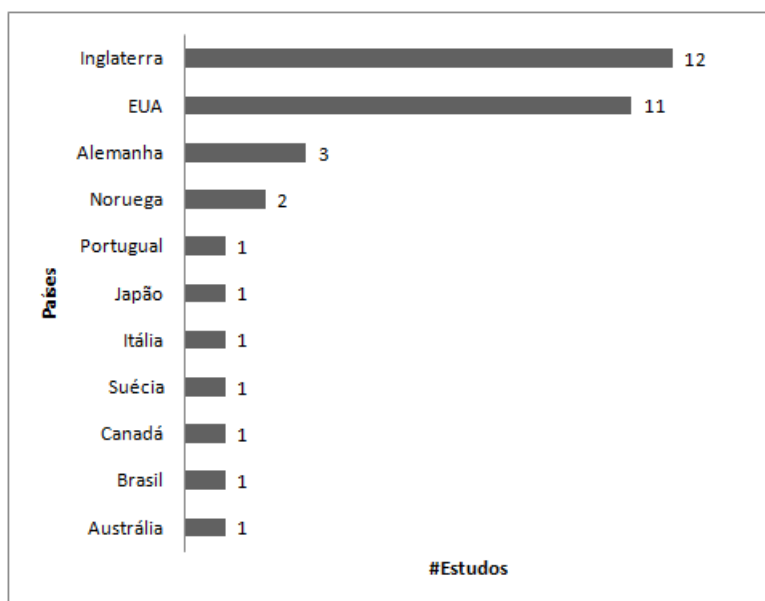


Figura 4.7 Distribuição dos estudos pelos países das instituições dos pesquisadores.

dividida em três partes, cada uma respondendo uma questão de pesquisa. A Seção 4.2.1 apresenta as evidências quanto aos tipos de mecanismos existentes para guiar estudos empíricos em ES. Na Seção 4.2.2 são apresentadas as evidências em relação a que tipo de estudos empíricos esses mecanismos guiam. E por fim, na Seção 4.2.3 são apresentadas as evidências de como os mecanismos encontrados estruturam o processo de condução dos estudos empíricos. Todas as evidências são apresentadas com a devida referência. Os estudos primários selecionados são referenciados com um prefixo EP mais uma numeração de 1 a 23 na ordem em que são dispostos na lista no Apêndice A.

4.2.1 Tipo de Mecanismo

Q1 - Quais são os mecanismos existentes para guiar estudos empíricos em engenharia de software?

O objetivo dessa questão é mapear os tipos de mecanismo usados para guiar estudos empíricos em engenharia de software. Nos 23 estudos analisados, foram encontrados 8 tipos de mecanismos que ajudam a guiar estudos empíricos; alguns estudos primários se enquadram em mais de um tipo. A Figura 4.8 ilustra, em ordem de frequência, os tipos de mecanismos identificados a partir dos estudos primários selecionados.

As evidências quanto ao tipo de mecanismo encontrado para guiar um estudo empírico

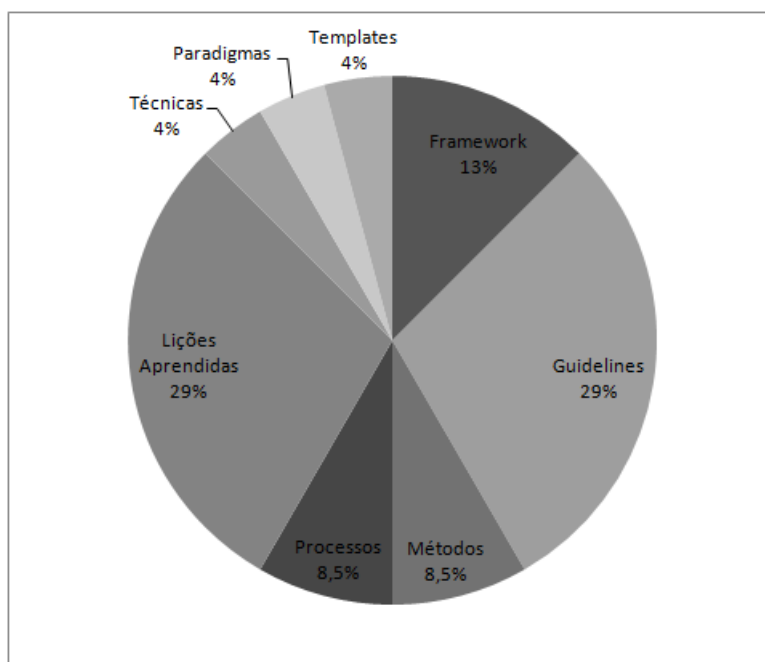


Figura 4.8 .

em ES estão resumidas na Tabela 4.2 e todas as transcrições coletadas como evidências são apresentadas ao longo desta seção.

Tabela 4.2 Resumo dos estudos por tipo de mecanismo encontrado.

Tipo de Mecanismo	Referências - EP: Estudos Primários	# de Estudos (%)
Framework	EP01, EP10, EP20	3 (13%)
Guideline	EP02, EP03, EP12, EP14, EP15, EP18, EP20	7 (29%)
Método	EP21, EP16	2 (8,5%)
Técnica	EP17	1 (4%)
Lições Aprendidas	EP04, EP05, EP06, EP07, EP08, EP09, EP11	7 (29%)
Processo	EP03, EP13	2 (8,5%)
Template	EP23	1 (4%)
Paradigma	EP19	1 (4%)

Como descrito no protocolo do mapeamento no Apêndice C, foi utilizado um código para marcar as passagens que foram utilizadas como amostras. Para esta questão, a cada novo tipo de mecanismo encontrado era associado um código novo, ao fim foi obtido:

- TM01: Framowork;
- TM02: Guideline;

- TM03: Lições Aprendidas;
- TM04: Processo;
- TM05: Técnica;
- TM06: Paradigma;
- TM07: Template;
- TM08: Método.

Frameworks

Um framework é um conjunto estruturado de conceitos ou soluções comuns a uma determinada área, cuja finalidade principal é obter reuso. Dessa forma ameniza parte do esforço de desenvolvimento de novas soluções. No que diz respeito a desenvolvimento de software, algumas definições podem ser citadas, uma delas é a de Johnson [52], onde um framework é um conjunto de classes que incorpora um *design* abstrato de soluções para um família de problemas relacionados. Outra definição é a de Givoni [53] que descreve um framework como uma coleção abstrata de classes, interfaces e padrões dedicados a solucionar uma classes de problemas através uma arquitetura flexível e extensível.

A seguir serão apresentados trechos coletados como evidências para classificar os estudos primários (EP01, EP10 e EP20) como frameworks.

- **EP01_TM01** - *“In this paper we lay out a **framework** for analyzing most of the experimental work that has been performed in software engineering over the past several years. The first objective is to describe a framework for experimentation in software engineering” ... “We have described a **framework** for experimentation to provide a structure for presenting previous studies. We also recommend the framework as a mechanism to facilitate the definition, planning, operation and interpretation of past and future studies.”*
- **EP10_TM01** - *“Our experience in conducting empirical software engineering from a variety of viewpoints for the last decade has made us aware of the criticality of integrating the various types of data that are collected and analyzed as well as the criticality of integrating the various types of activities that take place such as experiment design and the experiment itself. This has led us to develop a Computer-Aided Empirical Software Engineering (CAESE) **framework** as a substrate for*

*supporting the empirical software engineering lifecycle.” ... “We have presented Computer-Aided Empirical Software Engineering (CAESE) **framework** and the Ginger2 system, which is an instantiation of the CAESE framework.”*

- **EP20_TM01** - *“The first phase of the project has critically surveyed current evaluation approaches which were classified as ‘quantitative’ or ‘qualitative’. It was concluded that DESMET must support both types and provide guidelines to identify which method is appropriate to the particular circumstances of the user” ... “Although, DESMET proposes other courses of action in many circumstances, a special module has been designated in order to provide a new **framework** with practical guidelines for implementing formal experimentation, when this is appropriate, in software engineering environments.” ... “This paper proposes a **framework** for applying formal experimentation in the evaluation of software methods and tools.”*

Guidelines

Um guideline é algo, normalmente um documento, que visa melhorar ou orientar a execução de uma determinada rotina. Ele pode ser construído com base em dicas, conselhos, melhores práticas ou lições aprendidas de uma determinada área e pode ser intercambiado de uma área de estudo para outra quando este for adequado.

Os trechos coletados como evidências que classificou sete estudos primários (EP02, EP03, EP12, EP14, EP15, EP18 e EP22) como guidelines são apresentados a seguir.

- **EP02_TM02** - *“This paper argues that one can significantly reduce the risks associated with defining a new controlled experiment by running a set of well-planned observational pilot studies.” ... “The paper defines an approach to establish such pilot studies and illustrates it with a case study that shows the definition of an experiment to evaluate inspection techniques applied to data exploration tools through a set of four observational studies and one experimental trial. Based on the lessons learned we present some **guidelines** on how to test and improve experimental material in such way.”*
- **EP03_TM02** - *“First we want to arise the awareness of on-line surveys and discuss methods how to perform these in the context of software engineering. Second, we report our experience in performing on-line surveys in the form of lessons learned and **guidelines**.” ... “To guide the reader in performing on-line surveys we describe the process for conducting on-line surveys” ... “Our paper contributes by presenting a process of how to perform on-line surveys and by sharing our experience we gained in five on-line surveys in the form of **guidelines**.”*
- **EP12_TM02** - *“We propose a preliminary set of research **guidelines** aimed at stimulating discussion among software researchers.” ... “The **guidelines** are intended to assist researchers, reviewers and meta-analysts in designing, conducting and evaluating empirical studies.” ... “In this paper, however, we are concerned with developing **guidelines** to assist researchers to avoid major pitfalls in their research activities and to report their research correctly.”*
- **EP14_TM02** - *“The focus of this paper is on showing how qualitative methods can be adapted and incorporated into the designs of empirical studies in software engineering.” ... “The focus of this article has been to provide **guidance** on using qualitative research methods, particularly in studies in which they are combined with quantitative methods, in empirical studies of software engineering.”*

- **EP15_TM02** - *“Through a synthesis of existing best practices in case study research, we provide a set of comprehensive **guidelines** for conducting multiple case studies in software engineering research.”*
- **EP18_TM02** - *“This paper aims at providing an introduction to case study methodology and **guidelines** for researchers conducting case studies and readers studying reports of such studies.” ... “We present recommended practices for software engineering case studies as well as empirically derived and evaluated checklists for researchers and readers of case study research.”*
- **EP22_TM02** - *“This article provides **guidelines** for organizing and analyzing case studies so that they produce meaningful results.”*

Lições Aprendidas

Lições aprendidas é um conjunto de relatos sobre experiências relacionados geralmente a um determinado tema. Elas podem ser bastante úteis para que uma pessoa evite cometer erros que outras pessoas cometeram. São também de grande importância para identificar pontos de melhoria em uma determinada área.

Os trechos coletados como evidência para classificar os estudos primários EP04, EP05, EP06, EP07, EP08, EP09 e EP11 como lições aprendidas são:

- **EP04_TM03, EP05_TM03, EP06_TM03, EP07_TM03, EP08_TM03, EP09_TM03** - *“So this series of articles shares with you our **lessons learned**, in the hope of improving survey research in software engineering.”*
- **EP11_TM03** - *“...the focus of this paper is to convey some reflections and **lessons learned** in performing a large-scale, multinational survey of software engineering in the Information and Communication Technology (ICT) industry. ”*

Processos

Um processo é um conjunto de atividades parcialmente ordenadas (podem ser atividades concorrentes), cujo resultado é alcançar um objetivo. Um objetivo pode ser atingir uma meta, construir um produto ou gerar um artefato com valor agregado. Dependendo do contexto empregado o objetivo pode mudar. Em engenharia de software um processo é um conjunto de atividades e resultados associados que geram um produto de software [26].

Abaixo estão os trechos coletados como evidências que classificou dois estudos primários (EP03 e EP13) como processos.

- **EP03_TM04** - *“To guide the reader in performing on-line surveys we describe the **process** for conducting on-line surveys...” ... “Our paper contributes by presenting a **process** of how to perform on-line surveys and by sharing our experience we gained in five on-line surveys in the form of guidelines.”*
- **EP13_TM04** - *“In this paper, we model the experimental software engineering (ESE) **process**, capturing current experimental best practices.” ... “This paper proposed a model for the experimental software engineering **process**, using UML2 Activity Diagrams, that supports the definition and enactment of the activities, and corresponding deliverables, involved in ESE practice.”*

Técnicas

Uma técnica pode ser encarada como uma forma particular de fazer algo que pode ser representada por uma instrução ou um conjunto de instruções que visa concluir uma tarefa com o resultado desejado.

A seguir será apresentado o trecho coletado como evidência que classificou o estudo primário EP17 como uma técnica.

- **EP17_TM05** - *“Personal construct theory (applied via the repertory grid **technique**) supports interpretivist research in a structured manner and, as such, has relevance for researchers conducting studies focused on the human and organisational aspects of software engineering. Personal construct theory (which underpins the repertory grid **technique**) is introduced, and the technique and its administration is discussed.” ... “Here we address this situation for one research **technique**: repertory grids.”*

Paradigmas

Um paradigma é um modelo que determina como se atua sobre uma determinada realidade, é um conjunto de regras que estabelecem fronteiras, ou seja, o domínio dessa realidade e como resolver problemas dentro dela. Pode ser visto como um padrão de como resolver problemas em uma determinada área. Segundo Thomas Kuhn, um paradigma é o conjunto de práticas que definem uma disciplina científica em um período de tempo [54].

O trecho transcrito abaixo foi coletado como uma evidência que classificou o estudo primário EP19 como um paradigma.

- **EP19_TM06** - *“This paper outlines four **paradigms** for experimentation and empirical study in software engineering and describes their interrelationships: (1) Improvement paradigm (2) Goal-question-metric paradigm, (3) Experimentation framework paradigm, and (4) Classification paradigm.”*

Template

Os templates são modelos que ajudam a construção de algo através de uma estrutura (sem conteúdo) pré-construída. No caso desta estudo, um estudo primário foi identificado com o objetivo de fornecer um template que prover uma estrutura cujo objetivo é ajudar pesquisadores a construir um protocolo de estudo de caso. O objetivo do estudo é prover um mecanismo para melhorar o rigor na execução de estudos de caso em engenharia de software.

A seguir é apresentado o trecho coletado como evidência que classificou o estudo primário EP23 como um template.

- **EP23_TM07** - *“In order to undertake a series of case studies aimed at investigating systematic literature reviews, we have developed a case study protocol **template**. This paper introduces the **template** and discusses our experiences of using the **template** and the resulting case study protocol.” ... “In this paper we present a generic **template** that may help case study researchers to construct a case study protocol.” ... “This paper has introduced a **template** that can help the production of a case study protocol. We believe that this **template** will support Höst and Runeson (2007) checklists to improve the rigour of case studies in software engineering.”* A referência citada no trecho anteriormente transcrito pode ser vista em [55]

Método

Segundo Lakatos [49] um método é o conjunto de atividades sistemáticas e racionais que, com maior segurança e economia, permitem alcançar o objetivo - conhecimentos válidos e verdadeiros -, traçando o caminho a ser seguido, detectando erros e auxiliando as decisões do cientista. Na mesma linha segue Bunge [56] que define o método como sendo um procedimento regular, explícito e passível de ser repetido para conseguir-se alguma coisa, seja material ou conceitual.

A diante são apresentados os trechos coletados como evidências a partir dos estudos primários EP21 e EP16. Tais evidências foram responsáveis por classificar os estudos citados como métodos.

- **EP21_TM08** - *“This paper proposes dialog-based protocol for the study of the cognitive activities during software development and evolution.” ... “In this paper, we present a new empirical **method**, dialog-based protocol analysis...”*
- **EP16_TM08** - *“This paper demonstrated the usage of single-case designs and statistical tests for software engineering experiments with $N=1$. The principle and characteristics of single-case designs for software engineering experiments have been outlined.” ... “In conclusion we wish to stress that single-case studies can be a useful **method** in experimental software engineering, which has to be seen a supplement, not an alternative to other **methods**.”*

4.2.2 Tipo de Estudos Empíricos

Q2 - Os mecanismos encontrados são utilizados para guiar que tipo de estudo empírico?

Nesta questão abordaremos a questão Q2, que visa mapear os mecanismos encontrados para guiar estudos empíricos em ES por tipo de estudo. Os tipos de estudos empíricos mapeados por esse estudo são os descritos no Seção 2.3 do Capítulo 2

A seção é dividida por tipo de estudo empírico pesquisado. Action research, quasi-experimento e etnografia foram omitidas porque não apareceram nos estudos primários selecionados. O gráfico na Figura 4.9 ilustra a distribuição dos estudos primários selecionados por tipo de estudo empírico.

Como descrito no protocolo do mapeamento no Apêndice C, foi utilizado um código para marcar as passagens que foram utilizadas como amostras. Para esta questão, cada um dos tipos de estudos empíricos apresentados no Capítulo 2 recebeu um código, além destes a categoria Pesquisa Empírica foi criada e também recebeu um código, este termo foi utilizado para classificar guias que não tinham foco em nenhum estudo empírico específico.

- TEG01: Experimento Controlado;
- TEG02: Quasi-experimento;
- TEG03: Survey;
- TEG04: Estudo de Caso;
- TEG05: Pesquisa-ação;

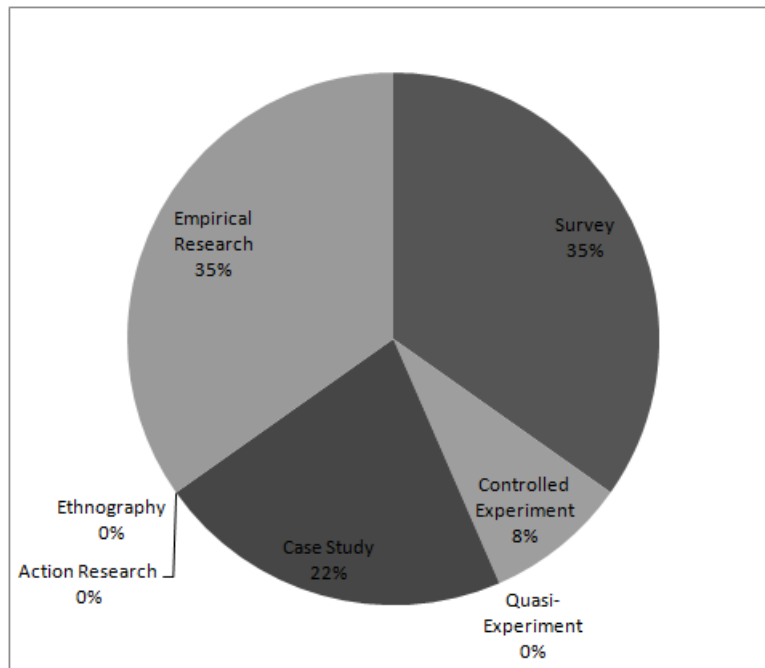


Figura 4.9 Distribuição dos EP selecionados por tipo de estudo abordado.

- TEG06: Etnografia;
- TEG07: Pesquisa Empírica.

Experimento Controlado

Abaixo são apresentadas as evidências coletadas a partir dos estudos primários EP02 e EP20 que classificaram o tipo de estudo empírico guiado por eles como experimentos controlados.

- **EP02_TEG01** - *“Controlled experiments are a key approach to evaluate and evolve our understanding of software engineering technologies.” ... “This paper argues that one can significantly reduce the risks associated with defining a new **controlled experiment** by running a set of well-planned observational pilot studies aimed at improving the experimental material.”*
- **EP20_TEG01** - *“Formal (Controlled) Experimentation has always been a powerful technique in the advance of traditional sciences;” ... “Although, DESMET proposes other courses of action in many circumstances, a special module has been designated in order to provide a new framework with practical guidelines*

*for implementing **formal experimentation**, when this is appropriate, in software engineering environments.”*

Survey

A seguir são apresentadas as evidências coletadas a partir dos estudos primários EP03, EP04, EP05, EP06, EP07, EP08, EP09 e EP11 que classificaram o tipo de estudo empírico guiado por eles como survey.

- **EP03_TEG03** - *“To guide the reader in performing on-line **surveys** we describe the process for conducting on-line **surveys** ...” ... “Our paper contributes by presenting a process of how to perform on-line **surveys** and by sharing our experience we gained in five on-line **surveys** in the form of guidelines.”*
- **EP04_TEG03, EP05_TEG03, EP06_TEG03, EP07_TEG03, EP08_TEG03, EP09_TEG03** - *“**Surveys** are probably the most commonly-used research method world-wide. **Survey** work is visible not only because we see many examples of it in software engineering research, but also because we are often asked to participate in **surveys** in our private capacity, as electors, consumers, or service users.” ... “This article and subsequent ones in this series describe how wrong we were. We do not want to give the impression that there is any way of turning a bad **survey** into a good one;” ... “So this series of articles shares with you our lessons learned, in the hope of improving **survey** research in software engineering.”*
- **EP11_TEG03** - *“... the focus of this paper is to convey some reflections and lessons learned in performing a large-scale, multinational **survey** of software engineering in the Information and Communication Technology (ICT) industry.”*

Estudo de Caso

Neste ponto são apresentadas as evidências coletadas a partir dos estudos primários EP15, EP16, EP18, EP22 e EP23 que classificaram o tipo de estudo empírico guiado por eles como estudo de caso.

- **EP15_TEG04** - *“Our detailed set of guidelines can help with the process of planning, constructing and reporting on all types of multiple **case studies**.” ... “Our guidelines are also appropriate for single **case study** design.”*

- **EP16_TEG04** - *“This approach is the use of **single-case studies** in combination with appropriate statistical methods” ... “This paper demonstrated the usage of single-case designs and statistical tests for software engineering experiments with $N=1$.”*
- **EP18_TEG04** - *“This paper aims at providing an introduction to **case study** methodology and guidelines for researchers conducting **case studies** and readers studying reports of such studies.” ... “We present recommended practices for software engineering **case studies** as well as empirically derived and evaluated checklists for researchers and readers of case study research.” ... “The purpose of this paper is to provide guidance for the researcher conducting **case studies**, for reviewers of **case study** manuscripts and for readers of **case study** papers.”*
- **EP22_TEG04** - *“This article provides guidelines for organizing and analyzing **case studies** so that they produce meaningful results.”*
- **EP23_TEG04** - *“In order to undertake a series of case studies aimed at investigating systematic literature reviews, we have developed a **case study** protocol template.” ... “In this paper we present a generic template that may help **case study** researchers to construct a **case study** protocol.” ... “This paper has introduced a template that can help the production of a **case study** protocol. We believe that this template will support Höst and Runeson (2007) checklists to improve the rigour of case studies in software engineering.”*

A referência citada no trecho anteriormente transcrito pode ser vista em [\[55\]](#)

Pesquisa Empírica

Abaixo são apresentadas as evidências coletadas dos estudos primários EP01, EP10, EP12, EP13, EP14, EP17, EP19 e EP21 que não estão relacionadas a nenhum tipo específico de estudo empírico. Desta estes estudos foram classificados como pesquisa empírica.

- **EP01_TEG07** - *“In this paper we lay out a framework for analyzing most of the **experimental work** that has been performed in software engineering over the past several years.”*
- **EP10_TEG07** - *“The recognition of these two integration needs has led us to develop the Computer-Aided Empirical Software Engineering (CAESE) framework as*

a substrate for supporting the **in vitro empirical software engineering** lifecycle.” ... “The second half of this paper then describes the Ginger2 environment as an instantiation of our framework. Ginger2 is an integrated environment that focuses on supporting **in vitro studies** in empirical software engineering.”

- **EP12_TEG07** - “The guidelines are intended to assist researchers, reviewers and meta-analysts in designing, conducting and evaluating **empirical studies**.” ... “In addition, empirical guidelines are often specialised to consider particular types of study e.g. randomised trials, surveys, exploratory studies. Clearly, the particular requirements for a set of guidelines influence their content and format. In the long term, if the software community accepts the need for experimental guidelines, we would expect to find specialised guidelines for different purposes. In this paper, however, we are concerned with developing guidelines to assist researchers to avoid major pitfalls in their research activities and to report their research correctly.”
- **EP13_TEG07** - “This process model can be used either as a support in the definition of **software engineering experiments** or in conducting comparisons among **experiment** results.” ... “In this section, we present a process model for **experimentation** in software engineering.”
- **EP14_TEG07** - “The focus of this paper is on showing how qualitative methods can be adapted and incorporated into the designs of **empirical studies** in software engineering.”
- **EP17_TEG07** - “Personal construct theory (applied via the repertory grid technique) supports **interpretivist research** in a structured manner and, as such, has relevance for researchers conducting studies focused on the human and organisational aspects of software engineering.” ... “The strength of personal construct theory and repertory grids in unearthing tacit knowledge and attitudes in a structured and manageable manner indicates that there are two key areas in which they can make a valuable contribution within **software engineering research**: exploration and evaluation.” ... “In exploratory situations (where the concern is to unearth tacit knowledge, information, preconceptions) qualitative approaches are generally required ” ... “In evaluative studies quantitative approaches are frequently adopted”
- **EP19_TEG07** - “The paradigms described in this paper are intended to help advance the use of measurement and empirical methods in software engineering. They offer a form of the scientific method for **experimentation** in the software domain.”

- **EP21_TEG07** - *“This paper propose a new empirical method that consists of dialog-based protocol, software screencapturing, and a coding scheme based on self-directed learning. The dialog-based protocol can greatly benefit software engineering **empirical research** by collecting more detailed and complete data, and reducing Hawthorne and placebo effects.”*

4.2.3 Estrutura dos Mecanismos

Q3 - Como se estruturam os mecanismos utilizados para guiar estudos empíricos?

A pergunta de pesquisa abordada neste seção visa descobrir como foram estruturados os mecanismos encontrados para guiar estudos empíricos em ES, ou seja, quais as etapas, fases ou elementos que compõem esses mecanismos. Com esse resultado é possível realizar uma análise para verificar semelhanças entre eles com o objetivo de criar um processo genérico que serviria de referência para área de ES ou até mesmo identificar lacunas entre eles com a finalidade de encontrar falhas e dessa forma propor melhorias.

Esta seção é dividida por tipo de estudo empírico pesquisado. Action research, quasi-experimento e etnografia foram omitidas porque não aparecem nos estudos primários selecionados.

Experimento Controlado

- **EP02** - *“The basis for our approach is the experimental paradigm for software engineering described by Wohlin in [19]. This paradigm consists of the following five phases: **Definition, Planning, Operation, Analysis, and Packaging**. We propose including a **testing phase** between the planning and operation phases, see Figure 1.”*

A figura referenciada no trecho anteriormente transcrito pode ser vista na Figura 4.10, bem como a referência em [1].

- **EP20** - *“This Framework rests on three main pillars. These are **Experimentation Viability, Experimentation Management and Experimentation Procedures**” ... “**Experiment Management**: It is advisable to appoint a team to overlook all technical, organisational and administrative aspects of the experimentation process. The main tasks of the team are:*
 - to **identify the viability** of formal experimentation,

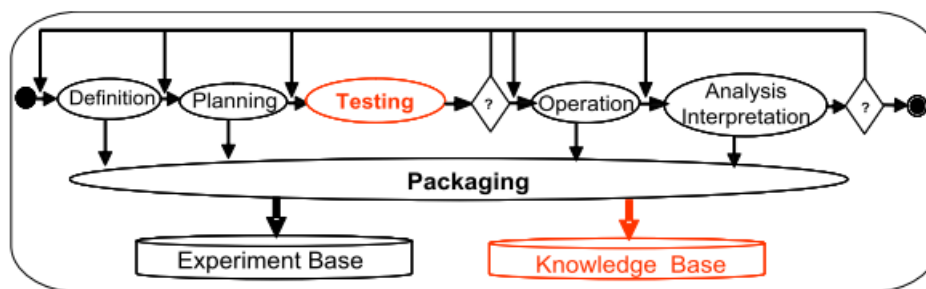


Figura 4.10 Experimentation Process [1].

- to *design the experiment*; and
- to *manage the implementation* of the experiment.”

...

“*Experimentation Procedures:*

- *Step One: Purpose Definition*
- *Step two: Focus the Scope*
- *Step three: State the Hypothesis*
- *Step four: Define The Treatment*
- *Step five: Identify Response Variables*
- *Step six: Select State Variables*
- *Step seven: Define Data Collection and Measurement Procedures*
- *Step eight: Choosing Experimental Objects*
- *Step nine: Choose Experimental Subjects”*

Survey

- **EP03** - “Our process to perform an on-line survey is an instantiation of a general process for conducting empirical studies in software engineering. Various instances of this process have been described meanwhile, see e.g., [6], [5], which is also applicable for the empirical strategies experiment and case study. For surveys the process is sketched in the following table.”

A tabela citada está disponível na Figura 4.11, bem como as referências podem ser vista em [57] e [9] respectivamente.

Activity	Purpose	Survey specific issues
Study definition	Determine the goal of the study	No specific issues
Study Design	Operationalize the study goals into a set of questions and select the respondents.	Questionnaire design, Define target population and sampling procedure Address validity issues.
Implementation	Operationalize the design so that the survey will be executable	Check completeness and understandability of questionnaire, Determine how to distribute the questionnaire (by post, mail), Define interview guidelines, Brief interviewers
Execution	Collect data and process data	Send out questionnaire, Collect filled-in questionnaires, Conduct interviews
Analysis	Interpret the data	Check on correct entering of data
Packaging	Report about the survey results	No specific issues

Figura 4.11 Activities in the process to perform on-line surveys [2].

- EP04, 05, 06, 07, 08, 09 - Estes estudos compõem uma série intitulada “*Principles of survey research*”
“It is a comprehensive system for collecting information to describe, compare or explain knowledge, attitudes and behavior. Thus, the survey instrument is part of a larger survey process with clearly-defined activities:
 1. *Setting specific, measurable objectives*
 2. *Planning and scheduling the survey*
 3. *Ensuring that appropriate resources are available*
 4. *Designing the survey*
 5. *Preparing the data collection instrument*
 6. *Validating the instrument*
 7. *Selecting participants*
 8. *Administering and scoring the instrument*
 9. *Analyzing the data*
 10. *Reporting the results”*
- EP11 - “*Section 3 discusses definition of population, and Section 4 how to make a sample from this population. Section 5 discusses questionnaire design, and Section 6 discusses the company contact process.*”

Estudo de Caso

- **EP15** - *“We organized our case study guidelines into several phases generally seen in case study research (but not necessarily documented in the case study literature). The phases are: **research initiation** or **pre-planning**, **administration**, **focus case study** or **planning**, **design case study plan**, **data collection**, **data analysis** (including evaluation and conclusions), and **reporting**.”*
- **EP16** - *“Experiments using a single-case experimental design ate characterized by several aspects compared with experiments that use factorial designs” ... “Especially, there are differences in **collecting data**, **establishing a base line**, **controlling carry-over effects**, **testing data dependency**, **testing change in level**, and in the **validity of results**.”*
- **EP18** - *“When conducting a case study, there are five major process steps to be walked through:*
 1. ***Case study design**: objectives are defined and the case study is planned.*
 2. ***Preparation for data collection**: procedures and protocols for data collection are defined.*
 3. ***Collecting evidence**: execution with data collection on the studied case.*
 4. ***Analysis of collected data**.*
 5. ***Reporting**”.*
- **EP22** - *“There are seven steps to follow in designing and administering case studies:*
 1. ***Define the hypothesis**.*
 2. ***Select the pilot projects**.*
 3. ***Identify the method of comparison**.*
 4. ***Minimize the effect of confounding factors**.*
 5. ***Plan the case study**.*
 6. ***Monitor the case study against the plan**.*
 7. ***Analyze and report the results**”.*
- **EP23** - *“... we based our template on Yin’s case study methodology while also incorporating some of Stake’s ideas. ... Using Yin’s approach, we produced a case*

study template suitable for our research program. We show a generalised version of the template in Appendix A. O Apêndice A deste trabalho é dividido da seguinte forma:

1. *Background ...*
2. *Design ...*
3. *Case Selection ...*
4. *Case Study Procedures and Roles ...*
5. *Data Collection ...*
6. *Analysis ...*
7. *Study Limitations ...*
8. *Plan Validity ...*
9. *Study Limitations ...*
10. *Reporting ...*
11. *Schedule ...*
12. *Appendices ...*”

Pesquisa Empírica

- **EP01** - *“The framework of experimentation, summarized in Fig. 1, consists of four categories corresponding to phases of the experimentation process: 1) **definition**, 2) **planning**, 3) **operation**, and 4) **interpretation**.”*

A figura referenciada no trecho anteriormente transcrito pode ser vista na Figura [4.12](#)

- **EP12** - *“We consider guidelines for what do to and what not to do under six basic topic areas: **Experimental context**, **Experimental design**, **Conduct of the experiment and Data collection**, **Analysis**, **Presentation of results**, **Interpretation of results**.”*

- **EP13** - *“Figure 1 presents an overview of the process. Its activities will be discussed in the next sub-sections.”*

A figura referenciada no trecho anteriormente transcrito pode ser vista na Figura [4.13](#)

I. Definition					
Motivation	Object	Purpose	Perspective	Domain	Scope
Understand	Product	Characterize	Developer	Programmer	Single project
Assess	Process	Evaluate	Modifier	Program/project	Multi-project
Manage	Model	Predict	Maintainer		Replicated project
Engineer	Metric	Motivate	Project manager		Blocked subject-project
Learn	Theory		Corporate manager		
Improve			Customer		
Validate			User		
Assure			Researcher		
II. Planning					
Design		Criteria		Measurement	
Experimental designs		Direct reflections of cost/quality		Metric definition	
Incomplete block		Cost		Goal-question-metric	
Completely randomized		Errors		Factor-criteria-metric	
Randomized block		Changes		Metric validation	
Fractional factorial		Reliability		Data collection	
Multivariate analysis		Correctness		Automatability	
Correlation		Indirect reflections of cost/quality		Form design and test	
Factor analysis		Data coupling		Objective vs. subjective	
Regression		Information visibility		Level of measurement	
Statistical models		Programmer comprehension		Nominal/classificatory	
Non-parametric		Execution coverage		Ordinal/ranking	
Sampling		Size		Interval	
		Complexity		Ratio	
III. Operation					
Preparation		Execution		Analysis	
Pilot study		Data collection		Quantitative vs. qualitative	
		Data validation		Preliminary data analysis	
				Plots and histograms	
				Model assumptions	
				Primary data analysis	
				Model application	
IV. Interpretation					
Interpretation context		Extrapolation		Impact	
Statistical framework		Sample representativeness		Visibility	
Study purpose				Replication	
Field of research				Application	

Figura 4.12 Sumário do framework para experimentação [3].

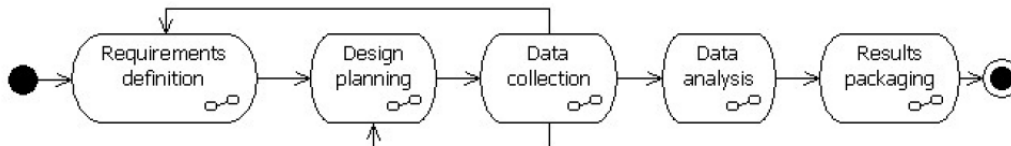


Figura 4.13 Visão geral do processo experimental [4].

- **EP14** - Este estudo primário discute uma série de métodos qualitativos sem que uma estrutura clara esteja disponível para ser extraída.
- **EP17** - “While the make up of a grid involves only three basic concepts (*elements, constructs and ratings scales*) there are a number of *design decisions* that need to be made during grid development that have the potential to affect both the reliability and validity of the grid and also the extent to which the technique can address the basic questions of the research.” ... “To ensure there is clarity about how and why repertory grids can be used in a research study we focus in this section

on the design decisions that need to be made in advance of their adoption.”

“The design decisions are divided as follows:

- 1. Element selection*
- 2. Elicitation of constructs*
- 3. The process of construct elicitation*
- 4. Rating elements and constructs*
- 5. Administration process*
- 6. Conduct the data collection process*
- 7. Analyse the data collected”*

- **EP19 - Improvement paradigm:** *“In its current form it has five essential aspects:*

- 1. Characterizing the environment ...*
- 2. Planning ...*
- 3. Execution ...*
- 4. Analysis ...*
- 5. Learning and feedback ...”*

Goal-question-metric paradigm: *“The **Goal/Question/Metric (GQM)** Paradigm is a mechanism for defining and evaluating a set of operational goals, using measurement on a specific project (see Fig. 1). It represents a systematic approach for setting the project **goals** tailored to the specific needs of an organization, defining them into a set of quantifiable **questions** that in turn implies a specific set of **metrics** and data collection. It involves the planning of the experimental framework. It includes the development of data collection mechanisms, eg. forms, automated tools, the collection and validation of data, and the analysis and interpretation of the collected data and computed metrics in the appropriate context of the questions and the original goals.”*

A figura referenciada no trecho anteriormente transcrito pode ser vista na Figura 4.14.

Experimentation framework paradigm: *“As defined in Ref. 18, it consists*

of four categories corresponding to phases of the experimentation process: (I) **definition**, (II) **planning**, (III) **operation**, and (IV) **interpretation**.”

Classification paradigm: “The three central activities in the paradigm are: (i) **data management and calibration**, (II) **classification tree generation**, and (III) **analysis and feedback** of newly acquired information to the current project.”

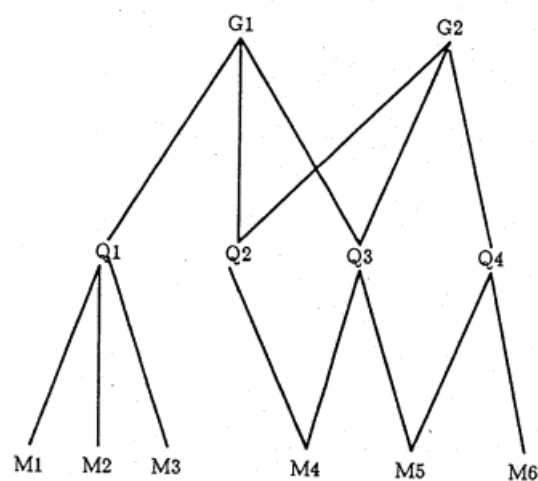


Figura 4.14 O paradigma Goal/Question/Metric (GQM). Goals= G_i , Questions= Q_i e Metrics= M_i [5].

- **EP21** - “In this paper, we present a new empirical method, dialog-based protocol analysis, see Figure 1.”

A figura referenciada no trecho anteriormente transcrito pode ser vista na Figura 4.15.

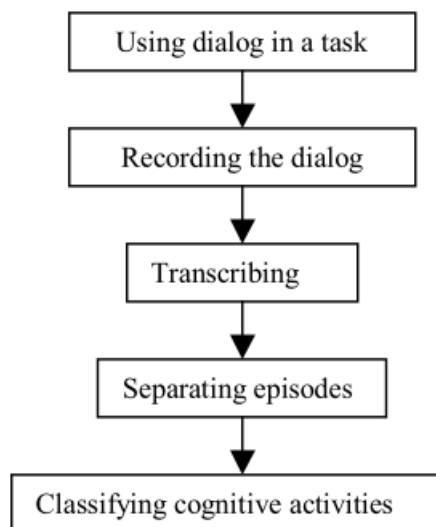


Figura 4.15 Passos do protocolo baseado em diálogo [6].

4.3 DISCUSSÃO SOBRE OS RESULTADOS

Apesar do pequeno número de trabalhos selecionados, o mapeamento respondeu satisfatoriamente as perguntas de pesquisa propostas.

A primeira pergunta de pesquisa visa descobrir o que foi desenvolvido especificamente para guiar estudos empíricos em ES. Respondendo a esta pergunta, o mapeamento identificou 8 tipos diferentes de mecanismos para guiar estudos empíricos, conforme mostrado na Figura 4.8. Um fato constatado a partir da análise deste resultado é que mais da metade, cerca de 58%, dos estudos primários estão concentrados em dois tipos de mecanismos: os guidelines e as lições aprendidas. Estes mecanismos tem origem na prática vivenciada pelos pesquisadores da área e, em alguns casos, em práticas de pesquisadores de outras áreas que possuem afinidade com a pesquisa empírica. No estudo EP12, por exemplo, pode ser observado um guideline desenvolvido para a comunidade de ES que foi baseado em experiências relatadas na área de pesquisa médica. Já no estudo EP11 pode ser observado o compartilhamento de experiências através de lições aprendidas sobre a execução de um *survey* em engenharia de software.

A segunda questão de pesquisa visa classificar os guias por tipo de estudo empírico. Um fato verificado no processo de responder esta pergunta foi que, conforme a Figura 4.9, três tipos de estudos empíricos não foram citados pelos guias selecionados, foram eles: quasi-experimento, pesquisa ação e etnografia. Algumas hipóteses são levantadas na

tentativa de explicar o ocorrido (confirmar tais hipóteses não é a intenção deste estudo):

- No caso de quasi-experimento pode ser explicado pelos estudos que foram selecionados para experimento controlado, já que eles podem ser usados para guiar quasi-experimentos devido à semelhança entre os dois tipos de estudos.
- Para pesquisa-ação e etnografia o que pode ser afirmado com este mapeamento é que não foram encontrados guias, específicos para ES, para estes dois tipos de estudos empíricos. No entanto, algumas hipóteses podem ser levantadas para explicar essa carência:
 - Podem ser estudos empíricos pouco executados em ES e, portanto, a comunidade científica ligada a ES ainda não desenvolveu guias específicos para eles. Isto poderia ser confirmado com uma pesquisa na literatura por estes tipos de estudos empíricos;
 - Outras áreas de pesquisa, como ciências médicas e sociais, podem ter mais experiência na execução destes tipos de estudos empíricos do que ES. Desta forma, os estudos conduzidos em ES podem se utilizar de guias oriundos de outras áreas de conhecimento. Como este mapeamento exige que os estudos sejam relacionados a ES, eventuais guias de outras áreas não compõem o conjunto de estudos primários selecionados. Isto pode ser confirmado através de uma pesquisa na literatura pelos trabalhos que realizaram estes dois tipos de estudo empírico em ES e verificar em que estudos os pesquisadores se basearam para planejar seus estudos;
 - Outra hipótese pode estar relacionada as limitações deste estudo, que pode ser encontrado no Capítulo 5.

Um resumo que apresenta a classificação dos estudos primários, de acordo com as respostas das questões Q1 e Q2, selecionados por este mapeamento por ser encontrado na Tabela 4.3.

A última questão, Q3, visa apresentar a estrutura que os mecanismos encontrados para guiar estudos empíricos propõem. A questão foi dividida por tipos de estudos empíricos. A partir da exposição destas estruturas, é possível desenvolver outras atividades interessantes para o desenvolvimento da engenharia de software empírica, tais como:

- Identificar problemas nos guias propostos através de uma comparação entre eles, ou seja, a partir de uma comparação realizada é possível identificar onde alguns guias podem estar em falta;

Tabela 4.3 Resumo dos resultados extraídos para questões Q1 e Q2

Código	Tipo de Mecanismo	Tipo de ES
EP01	Framework	Empirical Research
EP02	Guideline	Controlled Experiment
EP03	Process e Guideline	Survey
EP04	Lessons Learned	Survey
EP05	Lessons Learned	Survey
EP06	Lessons Learned	Survey
EP07	Lessons Learned	Survey
EP08	Lessons Learned	Survey
EP09	Lessons Learned	Survey
EP10	Framework	Empirical Research
EP11	Lessons Learned	Survey
EP12	Guideline	Empirical Research
EP13	Process	Empirical
EP14	Guideline	Empirical Research
EP15	Guideline	Case Study
EP16	Approach	Case Study
EP17	Technique	Empirical Research
EP18	Guideline	Case Study
EP19	Paradigms	Empirical Research
EP20	Framework e Guideline	Controlled Experiment
EP21	Method	Empirical Research
EP22	Guideline	Case Study
EP23	Template	Case Study

- Combinar as estruturas a fim de criar uma estrutura genérica que sirva de base para condução de estudos empíricos;
- A partir de uma estrutura genérica, que serviria de referência para uma ampla gama de estudos empíricos, poderiam ser derivadas estruturas para estudos específicos, que fossem mais robustas e confiáveis e que contemplem as peculiaridades de cada tipo de estudo empírico de forma que fossem seguidas pelos pesquisadores da comunidade de ES.

As uniões de várias características disponíveis nos estudos capturados por este mapeamento podem gerar novos guias que podem ser melhor estruturados. No entanto, tal guia só atingiria plenamente seu objetivo se usado extensivamente e rigorosamente pela comunidade a fim de gerar estudos empíricos mais homogêneos e dessa forma comparáveis. Sendo assim, seria possível construir conhecimento mais solidamente uma vez

que hipóteses poderiam ser apoiadas por vários estudos empíricos em conjunto.

Após observar as estruturas coletadas, é possível identificar uma semelhança entre elas. A maior parte dos guias tendem a seguir o modelo básico ilustrado na Figura 4.16.

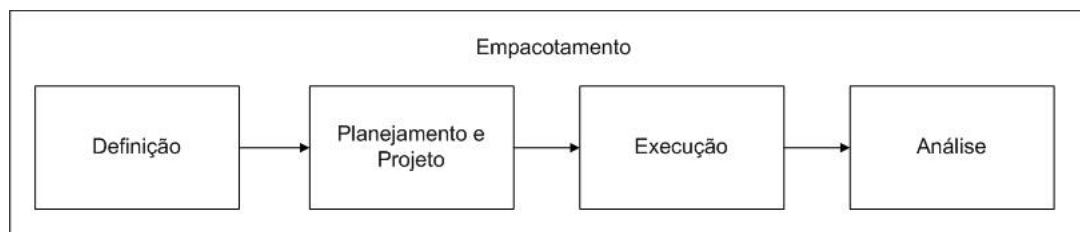


Figura 4.16 Estrutura base compartilhada entre maior parte dos guias.

Este trabalho não tem como objetivo desenvolver qualquer tipo de estrutura que vise ser uma referência, no momento. A estrutura apresentada acima é uma observação nos resultados da Q3. Uma descrição das etapas é apresentada mais detalhadamente abaixo:

- Definição: esta etapa está relacionada às atividades de definição do problema a ser investigado e o contexto no qual ele está inserido;
- Planejamento e Projeto: aglutina as atividades como geração das hipóteses do estudo, seleção das variáveis, seleção dos indivíduos que participarão do estudo, da instrumentação do estudo, definição dos procedimentos que serão utilizados e outras que por ventura colabore com o planejamento e projeto do estudo;
- Execução: os procedimentos planejados e projetados na etapa anterior são executados. Nesta etapa que são gerados os dados resultantes do estudo;
- Análise: está relacionada a atividades que interpretam os dados gerados na etapa anterior, análise estatística é um exemplo;
- Empacotamento: normalmente apresentada após as demais etapas, ela está relacionada a todas as outras. Ela é responsável pela geração do relatório que documentará todos os procedimentos utilizados e decisões tomadas durante a realização do estudo empírico.

O que pode ser afirmado após analisar os resultados apresentados pelo mapeamento é que há poucos mecanismos que guiam estudos empíricos especificamente desenvolvidos para ES. Não é possível afirmar que os pesquisadores não usem guias ou que não há guias disponíveis porque pode haver guias em outras áreas de pesquisa dentro e fora da

ciência da computação que estejam sendo usadas pelos pesquisadores e devido ao foco deste mapeamento, estes não foram incluídos.

4.4 RESUMO

Neste capítulo foram apresentados os resultados do mapeamento sistemático realizado. O processo de busca retornou um número total de 7101 estudos, dos quais 23 foram classificados como estudos primários relevantes. *IEEE* e *ACM* foram os que mais contribuíram para o mapeamento, com 7 estudos cada.

Os estudos primários selecionados apresentam oito mecanismos para guiar estudos empíricos, são eles: guidelines, lições aprendidas, frameworks, métodos, técnicas, paradigmas, templates e processos. Eles também foram agrupados por tipo de estudos empíricos que guiam. Dos cinco tipos de estudo empíricos apresentados, etnografia, pesquisa-ação e quasi-experimentos não tiveram nenhum estudo empírico encontram. São levantadas hipótese para explicar este fato. Foram coletadas as estruturas dos guias selecionados, e a partir delas são feitas observações, inclusive é possível visualizar uma estrutura básica seguida pela maioria dos estudos primários.

CONSIDERAÇÕES FINAIS

Neste capítulo, as considerações finais do estudo realizado são apresentadas. Entre elas são discutidas ameaças a validade do estudo, trabalhos futuros e as conclusões obtidas com o trabalho.

5.1 AMEAÇAS A VALIDADE

A primeira ameaça a validade diz respeito à estratégia de busca utilizada. A partir do momento que utilizamos engenhos de busca automatizados, estudos relevantes podem não ter sido inclusos no conjunto de estudos selecionados. Apesar de dedicarmos tempo para identificar palavras-chave relevantes, algum estudo em particular, que usa um termo diferente dos previstos, pode não ser identificado. Além disso, estudos recentes podem ainda não terem sido indexados pelos engenhos de busca. A busca manual realizada tenta minimizar esse risco, e cobriu os anais disponíveis de conferências relevantes para a área, tais como: *Empirical Software Engineering and Measurement (ESEM)* de 2007 a 2010 e *Evaluation and Assessment in Software Engineering (EASE)* de 2006 a 2010, e também foi feita uma busca manual no periódico *Empirical Software Engineering Journal (ESEJ)*, no período de 1997 a 2010.

Outro fator que pode ser encarado como uma ameaça é a pequena quantidade de estudos primários selecionados. Mesmo a busca automatizada sem qualquer restrição de período, apenas 23 estudos foram selecionados no final.

Por fim, de acordo com Kitchenham et al. [41], o fato dos dados deste trabalho terem sido extraídos por apenas um pesquisador configura uma limitação, mas é um fato aceitado por ela para alunos de PhD (não diz nada a respeito de alunos de mestrado). No entanto, a extração foi supervisionada pelo orientador do trabalho. Os demais pesquisadores que participaram do mapeamento, apresentados no Apêndice C, também contribuíram para sanar eventuais dúvidas que surgiram na etapa de extração e participaram ativamente da etapa de seleção dos estudos primários.

5.2 TRABALHOS FUTUROS

Um importante ponto que deve ser evidenciado em um trabalho é levantar as oportunidades de trabalhos futuros. Visando levantar essas oportunidades, a seguir são apresentadas algumas sugestões para novos encaminhamentos que foram identificados durante o estudo em questão.

- Aumentar a robustez do trabalho realizando uma avaliação de qualidade dos trabalhos selecionados. O objetivo é identificar se os guias propostos foram avaliados, e se é possível replicar a avaliação quando foram realizadas.
- Realizar uma busca através das referências bibliográficas declaradas nos estudos primários selecionados com a finalidade de selecionar novos estudos.
- Avaliar a completude dos guias propostos e verificar a possibilidade de integrar alguns deles com o objetivo de torná-los melhores e mais propícios a ser adotado como um modelo a ser seguido pela comunidade de engenharia de software.
- Verificar pontualmente os motivos pela falta de guias específicos para engenharia de software para pesquisa-ação e etnografia. E se for verificado realmente a falta destes, de onde é tirado embasamento teórico para a execução desses tipos de pesquisa em engenharia de software.
- E por fim, uma possibilidade é estender o protocolo a novos engenhos de busca e ampliar a busca manual a novas conferências e periódicos.

5.3 CONCLUSÕES

Este mapeamento sistemático de estudos procura na literatura por estudos que desenvolveram algum mecanismo de apoio a estudos empíricos em engenharia de software. Três perguntas de pesquisa foram usadas para guiar o estudo e após a avaliação de 7101 estudos, 23 estudos primários foram selecionados.

O processo de execução do mapeamento foi apresentado na dissertação permitindo a avaliação e a replicação do estudo por terceiros.

O *EI Compendex* se mostrou um engenho de busca interessante e promissor, sendo responsável pela maioria dos 7101 estudos retornados, isso indica que no futuro tal tipo de mecanismo pode ser o único utilizado pelo pesquisador, poupando tempo e esforço nessa etapa. A busca manual se mostrou interessante, com a contribuição de um estudo que

não foi capturado na busca automatizada. Além disso, a busca manual traz confiança ao pesquisador que percebe que estudos importantes que são encontrados na busca manual já foram identificados pela busca automatizada e captura estudos que usam termos que não foram previstos na construção da *string* de busca.

No entanto, as duas principais contribuições deste estudo estão primeiramente na estruturação e resumo de forma sistemática dos principais mecanismos para guiar estudos empíricos em engenharia de software, o que facilitará a condução de novos estudos empíricos, já que a indicação de algumas formas de fazê-los pode ser encontrada nesta dissertação. Além disso, a apresentação de lacunas na área também serve como uma indicação de trabalhos futuros para uma pesquisa mais profunda com o intuito de se confirmar a lacuna ou principalmente, de preenchê-la.

APÊNDICE A

ESTUDOS PRIMÁRIOS

ID	Ano	Fonte	Referência
EP01	1986	Scopus	V R Basili, R W Selby, and D H Hutchens. <i>Experimentation in software engineering</i> . IEEE Trans. Softw. Eng., 12(7):733-743, 1986.
EP02	2006	ACM	Manoel Mendonça, Daniela Cruzes, Josemeire Dias, and Maria Cristina Ferreira de Oliveira. <i>Using observational pilot studies to test and improve lab packages</i> . In Proceedings of the ACM/IEEE international symposium on Empirical software engineering, ISESE '06, pages 48-57, New York, NY, USA, 2006. ACM.
EP03	2003	IEEE	Teade Punter, Marcus Ciolkowski, Bernd Freimut, and Isabel John. <i>Conducting on-line surveys in software engineering</i> . In Proceedings of the International Symposium on Empirical Software Engineering, ISESE '03, page 80, Washington, DC, USA, 2003. IEEE Computer Society.
EP04	2001	ACM	Shari Lawrence Pfleeger and Barbara A. Kitchenham. <i>Principles of survey research: part 1: turning lemons into lemonade</i> . SIGSOFT Softw. Eng. Notes, 26(6):16-18, November 2001.
EP05	2002	ACM	Barbara A. Kitchenham and Shari Lawrence Pfleeger. <i>Principles of survey research part 2: designing a survey</i> . SIGSOFT Softw. Eng. Notes, 27(1):18-20, January 2002.
EP06	2002	ACM	Barbara A. Kitchenham and Shari Lawrence Pfleeger. <i>Principles of survey research part 3: constructing a survey instrument</i> . SIGSOFT Softw. Eng. Notes, 27(2):20-24, March 2002.
EP07	2002	ACM	Barbara Kitchenham and Shari Lawrence Pfleeger. <i>Principles of survey research part 4: questionnaire evaluation</i> . SIGSOFT Softw. Eng. Notes, 27(3):20-23, May 2002.
EP08	2002	ACM	Barbara Kitchenham and Shari Lawrence Pfleeger. <i>Principles of survey research part 5: populations and samples</i> . SIGSOFT Softw. Eng. Notes, 27(5):17-20, September 2002.
EP09	2003	ACM	Barbara Kitchenham and Shari Lawrence Pfleeger. <i>Principles of survey research part 6: data analysis</i> . SIGSOFT Softw. Eng. Notes, 28(2):24-27, March 2003.
EP10	1999	IEEE	Koji Torii, Ken-ichi Matsumoto, Kumiyo Nakakoji, Yoshihiro Takada, Shingo Takada, and Kazuyuki Shima. <i>Ginger2: An environment for computer-aided empirical software engineering</i> . IEEE Trans. Softw. Eng., 25:474-492, July 1999.
EP11	2005	IEEE	R. Conradi, J. Li, O.P.N. Slyngstad, V.B. Kampenes, C. Bunse, M. Morisio, and M. Torchiano. <i>Reflections on conducting an international survey of software engineering</i> . In International Symposium on Empirical Software Engineering, page 10 pp., 2005.

ID	Ano	Fonte	Referência
EP12	2002	IEEE	Barbara A. Kitchenham, Shari Lawrence Pfleeger, Lesley M. Pickard, Peter W. Jones, David C. Hoaglin, Khaled El Emam, and Jarrett Rosenberg. <i>Preliminary guidelines for empirical research in software engineering</i> . IEEE Trans. Softw. Eng., 28:721-734, August 2002.
EP13	2007	IEEE	Miguel Goulao and Fernando Brito e. Abreu. <i>Modeling the experimental software engineering process</i> . In Proceedings of the 6th International Conference on Quality of Information and Communications Technology, pages 77-90, Washington, DC, USA, 2007. IEEE Computer Society.
EP14	1999	IEEE	Carolyn B. Seaman. <i>Qualitative methods in empirical studies of software engineering</i> . IEEE Trans. Softw. Eng., 25(4):557-572, 1999.
EP15	2009	IEEE	J.M. Verner, J. Sampson, V. Tomic, N.A.A. Bakar, and B.A. Kitchenham. <i>Guidelines for industrially-based multiple case studies in software engineering</i> . In Third International Conference on Research Challenges in Information Science (RCIS)., pages 313 -324, april 2009.
EP16	2001	Science Direct	Andreas Zendler, Erika Horn, Heinz Schwärtzel, and Erhard Plöde-reder. <i>Demonstrating the usage of single-case designs in experimental software engineering</i> . Information & Software Technology, 43(12):681-691, 2001.
EP17	2009	Science Direct	Helen M. Edwards, Sharon McDonald, and S. Michelle Young. <i>The repertory grid technique: Its place in empirical software engineering research</i> . Inf. Softw. Technol., 51(4):785-798, April 2009.
EP18	2009	EI Compendex	Per Runeson and Martin Höst. <i>Guidelines for conducting and reporting case study research in software engineering</i> . Empirical Softw. Engg. Journal, 14:131-164, April 2009.
EP19	1991	EI Compendex	Victor R. Basili and Richard W. Selby. <i>Paradigms for experimentation and empirical studies in software engineering</i> . Reliability Engineering and System Safety, 32(1-2):171-191, 1991.
EP20	1993	EI Compendex	W E Mohamed, C J Sadler, and D Law. <i>Experimentation in software engineering: A new framework</i> . In Proceedings of Software Quality Management, Elsevier Science and CMP, Southampton, pages 417-430, 1993.
EP21	2005	EI Compendex	null Shaochun Xu and V. Rajlich. <i>Dialog-based protocol: an empirical research method for cognitive activities in software engineering</i> . Empirical Software Engineering, International Symposium on, 0:10, 2005
EP22	1995	Scopus	Barbara Kitchenham, Lesley Pickard, and Shari Lawrence Pfleeger. <i>Case studies for method and tool evaluation</i> . IEEE Softw., 12:52-62, July 1995.
EP23	2008	EASE	Pearl Brereton, Barbara Kitchenham, David Budgen, and Zhi Li. <i>Using a protocol template for case study planning</i> . In 12th International Conference on Evaluation and Assessment in Software Engineering (EASE), 2008.

APÊNDICE B

ESTUDOS EXCLUÍDOS

ID	Ano	Fonte	Referência	Critério
1	2008	IEEE	Darja Smite, Claes Wohlin, Robert Feldt, and Tony Gorschek. <i>Reporting empirical research in global software engineering: A classification scheme</i> . In Proceedings of the IEEE International Conference on Global Software Engineering, pages 173-181, Washington, DC, USA, 2008. IEEE Computer Society	Irrelevante
2	2007	IEEE	Martin Host and Per Runeson. <i>Checklists for software engineering case study research</i> . In Proceedings of the First International Symposium on Empirical Software Engineering and Measurement, ESEM '07, pages 479-481, Washington, DC, USA, 2007. IEEE Computer Society.	Incompleto
3	2004	IEEE	Andreas Jedlitschka and Marcus Ciolkowski. <i>Towards evidence in software engineering</i> . In Proceedings of the International Symposium on Empirical Software Engineering, pages 261-270, Washington, DC, USA, 2004. IEEE Computer Society.	Irrelevante
4	1999	IEEE	Victor R. Basili, Forrest Shull, Filippo Lanubile, and IEEE Computer Society. <i>Building knowledge through families of experiments</i> . IEEE Transactions on Software Engineering, 25:456-473, 1999.	Irrelevante*
5	1995	IEEE	J.A. Hughes. <i>Ethnography, plans and software engineering</i> . In CSCW (Computer Supported Cooperative Working) and the Software Process (Digest No. 1995/036), IEEE Colloquium on, pages 1-8, feb 1995.	Irrelevante
6	2008	IEEE	Guilherme H. Travassos, Paulo Sérgio Medeiros dos Santos, Paula Gomes Mian, Arilo Cláudio Dias Neto, and Jorge Biolchini. <i>An environment to support large scale experimentation in software engineering</i> . In Proceedings of the 13th IEEE International Conference on Engineering of Complex Computer Systems, pages 193-202, Washington, DC, USA, 2008. IEEE Computer Society.	Irrelevante
7	2007	IEEE	Dag I. K. Sjoberg, Tore Dyba, and Magne Jorgensen. <i>The future of empirical methods in software engineering research</i> . In Future of Software Engineering, FOSE '07, pages 358-378, Washington, DC, USA, 2007. IEEE Computer Society.	Irrelevante

ID	Ano	Fonte	Referência	Critério
8	2007	IEEE	Walter F. Tichy and Frank Padberg. <i>Empirical methods in software engineering research</i> . In International Conference on Software Engineering Companion, pages 163-164, Los Alamitos, CA, USA, 2007. IEEE Computer Society.	Incompleto
9	1996	IEEE	Victor R. Basili. <i>The role of experimentation in software engineering: past, current, and future</i> . In Proceedings of the 18th international conference on Software engineering, ICSE '96, pages 442-449, Washington, DC, USA, 1996. IEEE Computer Society.	Irrelevante
10	2006	IEEE	Victor R. Basili. <i>The role of empirical study in software engineering</i> . In Proceedings of the 30th Annual IEEE/NASA Software Engineering Workshop, pages 3-6, Washington, DC, USA, 2006. IEEE Computer Society.	Incompleto
11	1980	IEEE	B. Curtis. <i>Measurement and experimentation in software engineering</i> . Proceedings of the IEEE, 68(9):1144 - 1157, september 1980.	Irrelevante
12	1998	IEEE	Marvin V. Zelkowitz and Dolores R. Wallace. <i>Experimental models for validating technology</i> . Computer, 31:23-31, May 1998.	Irrelevante*
13	2007	ACM	Walter F. Tichy and Frank Padberg. <i>Empirical methods in software engineering research</i> . In International Conference on Software Engineering Companion, pages 163-164, Los Alamitos, CA, USA, 2007. IEEE Computer Society.	Estudo repetido, 1ª fonte: IEEE
14	2006	ACM	Barbara Kitchenham, Hiyam Al-Khilidar, Muhammad Ali Babar, Mike Berry, Karl Cox, Jacky Keung, Felicia Kurniawati, Mark Staples, He Zhang, and Liming Zhu. <i>Evaluating guidelines for empirical software engineering studies</i> . In Proceedings of the ACM/IEEE international symposium on Empirical software engineering, ISESE '06, pages 38-47, New York, NY, USA, 2006. ACM.	Irrelevante
15	2007	ACM	Steve Easterbrook. <i>Empirical research methods for software engineering</i> . In Proceedings of the twenty-second IEEE/ACM international conference on Automated software engineering, ASE '07, pages 574-574, New York, NY, USA, 2007. ACM.	Incompleto
16	1981	ACM	Thomas Moher and G. Michael Schneider. <i>Methods for improving controlled experimentation in software engineering</i> . In Proceedings of the 5th international conference on Software engineering, ICSE '81, pages 224-233, Piscataway, NJ, USA, 1981. IEEE Press.	Irrelevante
17	2007	ACM	Dag I. K. Sjøberg, Tore Dyba, and Magne Jørgensen. <i>The future of empirical methods in software engineering research</i> . In Future of Software Engineering, FOSE '07, pages 358-378, Washington, DC, USA, 2007. IEEE Computer Society.	Estudo repetido, 1ª fonte: IEEE

ID	Ano	Fonte	Referência	Critério
18	2009	ACM	Gregory Gay, Tim Menzies, Bojan Cukic, and Burak Turhan. <i>How to build repeatable experiments</i> . In Proceedings of the 5th International Conference on Predictor Models in Software Engineering, PROMISE '09, pages 15:1-15:9, New York, NY, USA, 2009. ACM.	Irrelevante
19	2010	ACM	Helen Sharp, Cleidson deSouza, and Yvonne Dittrich. <i>Using ethnographic methods in software engineering research</i> . In Proceedings of the 32nd ACM/IEEE International Conference on Software Engineering - Volume 2, ICSE '10, pages 491-492, New York, NY, USA, 2010. ACM.	Incompleto
20	2005	ACM	WenQian Liu, Charles L. Chen, Vidya Lakshminarayanan, and Dewayne E. Perry. <i>A design for evidence - based soft research</i> . In Proceedings of the 2005 workshop on Realising evidence-based software engineering, REBSE '05, pages 1-7, New York, NY, USA, 2005. ACM.	Irrelevante
21	2010	ACM	Roel Wieringa. <i>Design science methodology: principles and practice</i> . In Proceedings of the 32nd ACM/IEEE International Conference on Software Engineering - Volume 2, ICSE '10, pages 493-494, New York, NY, USA, 2010. ACM.	Incompleto
22	2003	ACM	Teade Punter, Marcus Ciolkowski, Bernd Freimut, and Isabel John. <i>Conductin gon-line surveys in software engineering</i> . In Proceedings of the 2003 International Symposium on Empirical Software Engineering, ISESE '03, page 80, Washington, DC, USA, 2003. IEEE Computer Society.	Estudo repetido, 1ª fonte: IEEE
23	1982	Science Direct	Thomas Moher and G. Michael Schneider. <i>Methodology and experimental research in software engineering</i> . International Journal of Man-Machine Studies, 16(1):65-87, 1982.	Artigo não encontrado
24	2008	Science Direct	Lorin Hochstein, Taiga Nakamura, Forrest Shull, Nico Zazworka, Victor R. Basili, and Marvin V. Zelkowitz. <i>Chapter 5 an environment for conducting families of software engineering experiments</i> . In Marvin V. Zelkowitz, editor, Software Development, volume 74 of Advances in Computers, pages 175-200. Elsevier, 2008.	Irrelevante
25	1982	Science Direct	Marvin V. Zelkowitz. <i>Data collection and evaluation for experimental computer science research</i> . In Proceedings of the Empirical Foundations for Computer and Information Science, pages 269-276, Atlanta, GA, 1982.	Irrelevante

ID	Ano	Fonte	Referência	Critério
26	2006	Scopus	Barbara Kitchenham, Hiyam Al-Khilidar, Muhammad Ali Babar, Mike Berry, Karl Cox, Jacky Keung, Felicia Kurniawati, Mark Staples, He Zhang, and Liming Zhu. <i>Evaluating guidelines for empirical software engineering studies</i> . In Proceedings of the ACM/IEEE international symposium on Empirical software engineering, ISESE '06, pages 38-47, New York, NY, USA, 2006. ACM.	Estudo repetido, 1 ^a fonte: ACM
27	2008	Scopus	Barbara Kitchenham, Hiyam Al-Khilidar, Muhammed Ali Babar, Mike Berry, Karl Cox, Jacky Keung, Felicia Kurniawati, Mark Staples, He Zhang, and Liming Zhu. <i>Evaluating guidelines for reporting empirical software engineering studies</i> . Empirical Softw. Eng., 13:97-121, February 2008.	Irrelevante
28	2007	Scopus	Steve Easterbrook. <i>Empirical research methods for software engineering</i> . In Proceedings of the twenty-second IEEE/ACM international conference on Automated software engineering, ASE '07, pages 574-574, New York, NY, USA, 2007. ACM.	Estudo repetido, 1 ^a fonte: ACM
29	2003	Scopus	Claes Wohlin, Martin H"ost, and Kennet Henningsson. <i>Empirical research methods in software engineering</i> . In Reidar Conradi and Alf Inge Wang, editors, ESERNET, volume 2765/2003 of Lecture Notes in Computer Science, pages 7-23. Springer, 2003.	Incompleto
30	2007	Scopus	Walter F. Tichy and Frank Padberg. <i>Empirical methods in software engineering research</i> . In International Conference on Software Engineering Companion, pages 163-164, Los Alamitos, CA, USA, 2007. IEEE Computer Society.	Estudo repetido, 1 ^a fonte: IEEE
31	2006	Scopus	Barbara Kitchenham. <i>Empirical paradigm - the role of experiments</i> . In Proceedings of the international conference on Empirical software engineering issues: critical assessment and future directions, pages 25-32, Berlin, Heidelberg, 2007. Springer-Verlag.	Irrelevante
32	2007	Scopus	James Miller. <i>Creating real value in software engineering experiments</i> . In Proceedings of the international conference on Empirical software engineering issues: critical assessment and future directions, pages 38-38, Dagstuhl Castle, Germany, 2006. Springer-Verlag.	Incompleto
33	2006	Scopus	Manoel Mendonça, Daniela Cruzes, Josemeire Dias and Maria Cristina Ferreira de Oliveira. <i>Using observational pilot studies to test and improve lab packages</i> . In Proceedings of the ACM/IEEE international symposium on Empirical software engineering, ISESE '06, pages 48-57, New York, NY, USA, 2006. ACM.	Estudo repetido, 1 ^a fonte: ACM
34	2006	Scopus	Markku Oivo. <i>New opportunities for empirical research</i> . In Proceedings of the 2006 international conference on Empirical software engineering issues: critical assessment and future directions, pages 22-22, Berlin, Heidelberg, 2007. Springer-Verlag.	Incompleto

ID	Ano	Fonte	Referência	Critério
35	1998	Scopus	Babak Khazaei, Jawed I. A. Siddiqi, Rick Osborn, and Chris Roast. <i>Templates and tariffs - tools for empirical studies</i> . Information and Software Technology, 40(14):823-829, 1998.	Irrelevante
36	2007	EI compendex	Steve Easterbrook. <i>Empirical research methods for software engineering</i> . In Proceedings of the twenty-second IEEE/ACM international conference on Automated software engineering, ASE '07, pages 574-574, New York, NY, USA, 2007. ACM.	Estudo repetido, 1ª fonte: ACM
37	2009	EI compendex	Per Runeson and Martin Höst. <i>Tutorial: Case studies in software engineering</i> . In Will Aalst, John Mylopoulos, Norman M. Sadeh, Michael J. Shaw, Clemens Szyperski, Frank Bomarius, Markku Oivo, Päivi Jaring, and Pekka Abrahamsson, editors, Product-Focused Software Process Improvement, volume 32 of Lecture Notes in Business Information Processing, pages 441-442. Springer Berlin Heidelberg, 2009.	Incompleto
38	1985	EI compendex	V R Basili, R W Selby, and D H Hutchens. <i>Experimentation in software engineering</i> . IEEE Trans. Softw. Eng., 12(7):733-743, 1986.	Estudo repetido, 1ª fonte:Scopus
39	2008	EI compendex	Barbara Kitchenham, Hiyam Al-Khilidar, Muhammed Ali Babar, Mike Berry, Karl Cox, Jacky Keung, Felicia Kurniawati, Mark Staples, He Zhang, and Liming Zhu. <i>Evaluating guidelines for reporting empirical software engineering studies</i> . Empirical Softw. Eng., 13:97-121, February 2008.	Estudo repetido, 1ª fonte: Scopus
40	1999	EI compendex	Koji Torii, Ken-ichi Matsumoto, Kumiyo Nakakoji, Yoshihiro Takada, Shingo Takada, and Kazuyuki Shima. <i>Ginger2: An environment for computer-aided empirical software engineering</i> . IEEE Trans. Softw. Eng., 25:474-492, July 1999.	Estudo repetido, 1ª fonte: IEEE
41	2008	EI compendex	Guilherme H. Travassos, Paulo Sérgio Medeiros dos Santos, Paula Gomes Mian, Arilo Cláudio Dias Neto, and Jorge Biolchini. <i>An environment to support large scale experimentation in software engineering</i> . In Proceedings of the 13th IEEE International Conference on on Engineering of Complex Computer Systems, pages 193-202, Washington, DC, USA, 2008. IEEE Computer Society.	Estudo repetido, 1ª fonte: IEEE
42	2005	EI compendex	A. Jedlitschka and D. Pfahl. <i>Reporting guidelines for controlled experiments in software engineering</i> . In Empirical Software Engineering International Symposium on, page 10, Los Alamitos, CA, USA, november 2005.	Irrelevante*

ID	Ano	Fonte	Referência	Critério
43	2009	EI compendex	Vladimir Mandic, Jouni Markkula, and Markku Oivo. <i>Towards multi-method research approach in empirical software engineering</i> . In Will Aalst, John Mylopoulos, Norman M. Sadeh, Michael J. Shaw, Clemens Szyperski, Frank Bomarius, Markku Oivo, Päivi Jaring, and Pekka Abrahamsson, editors, <i>Product-Focused Software Process Improvement</i> , volume 32 of <i>Lecture Notes in Business Information Processing</i> , pages 96-110. Springer Berlin Heidelberg, 2009.	Artigo não encontrado
44	2007	EI compendex	Martin Host and Per Runeson. <i>Checklists for software engineering case study research</i> . In <i>Proceedings of the First International Symposium on Empirical Software Engineering and Measurement, ESEM '07</i> , pages 479-481, Washington, DC, USA, 2007. IEEE Computer Society.	Estudo repetido, 1ª fonte: IEEE
45	1994	EI compendex	C. Seaman and L.G. Votta. <i>Session 4: Experimentation in the process framework</i> . In <i>Proceedings of Software Process Workshop, Ninth International</i> , page 15, october 1994.	Irrelevante
46	2000	EI compendex	Linden J. Ball and Thomas C. Ormerod. <i>Putting ethnography to work: the case for a cognitive ethnography of design</i> . <i>Int. J. Hum.-Comput. Stud.</i> , 53:147-168, July 2000.	Irrelevante*
47	1997	EI compendex	Jesper Simonsen and Finn Kensing. <i>Using ethnography in contextual design</i> . <i>Communications of the ACM</i> , 40(7):82-88, 1997.	Irrelevante*
48	2005	EI compendex	Hyunsook Do, Sebastian Elbaum, and Gregg Rothermel. <i>Supporting controlled experimentation with testing techniques: An infrastructure and its potential impact</i> . <i>Empirical Softw. Engg.</i> , 10(4):405-435, October 2005.	Irrelevante*
49	1996	EI compendex	Victor R. Basili. <i>The role of experimentation in software engineering: past, current, and future</i> . In <i>Proceedings of the 18th international conference on Software engineering, ICSE '96</i> , pages 442-449, Washington, DC, USA, 1996. IEEE Computer Society.	Estudo repetido, 1ª fonte: IEEE
50	1995	EI compendex	Lawrence G. Votta and Adam Porter. <i>Experimental software engineering: a report on the state of the art</i> . In <i>Proceedings of the 17th international conference on Software engineering, ICSE '95</i> , pages 277-279, New York, NY, USA, 1995. ACM.	Irrelevante*

ID	Ano	Fonte	Referência	Critério
51	2008	EI compendex	Yvonne Dittrich, Kari Rönkkö, Jeanette Eriksson, Christina Hansson, and Olle Lindeberg. <i>Cooperative method development. combining qualitative empirical research with method, technique and process improvement</i> . Empirical Software Engineering, 13:231-260, 2008.	Irrelevante*
52	2002	EI compendex	Jerry B. Weinberg and Mary L. Stephen. <i>Participatory design in a human-computer interaction course: teaching ethnography methods to computer scientists</i> . SIGCSE Bull., 34:237-241, February 2002.	Irrelevante *
53	1986	EI compendex	Lyn Antill. <i>Action research in information systems design</i> . Design Studies, 7(4):192-198, 1986. Special Issue: the design coalition team.	Artigo não encontrado
54	2004	EI compendex	Forrest Shull, Manoel G. Mendonça, Victor Basili, Jeffrey Carver, José C. Maldonado, Sandra Fabbri, Guilherme Horta Travassos, and Maria Cristina Ferreira. <i>Knowledge-sharing issues in experimental software engineering</i> . Empirical Softw. Engg., 9:111-137, March 2004.	Irrelevante
55	2006	EI compendex	Markku Oivo. <i>New opportunities for empirical research</i> . In Proceedings of the 2006 international conference on Empirical software engineering issues: critical assessment and future directions, pages 22-22, Berlin, Heidelberg, 2007. Springer-Verlag.	Estudo repetido, 1ª fonte: Scopus
56	1982	EI compendex	Thomas Moher and G. Michael Schneider. <i>Methodology and experimental research in software engineering</i> . International Journal of Man-Machine Studies, 16(1):65-87, 1982.	Estudo repetido, 1ª fonte: Science Direct
57	2006	EI compendex	Barbara Kitchenham, Hiyam Al-Khilidar, Muhammad Ali Babar, Mike Berry, Karl Cox, Jacky Keung, Felicia Kurniawati, Mark Staples, He Zhang, and Liming Zhu. <i>Evaluating guidelines for empirical software engineering studies</i> . In Proceedings of the ACM/IEEE international symposium on Empirical software engineering, ISESE '06, pages 38-47, New York, NY, USA, 2006. ACM.	Estudo repetido, 1ª fonte: ACM
58	2010	EI compendex	Helen Sharp, Cleidson deSouza, and Yvonne Dittrich. <i>Using ethnographic methods in software engineering research</i> . In Proceedings of the 32nd ACM/IEEE International Conference on Software Engineering - Volume 2, ICSE '10, pages 491-492, New York, NY, USA, 2010. ACM.	Estudo repetido, 1ª fonte: ACM

ID	Ano	Fonte	Referência	Critério
59	2008	EI compendex	Manoel G. Mendonça, José C. Maldonado, Maria C. F. de Oliveira, Jeffrey Carver, Sandra C. P. F. Fabbri, Forrest Shull, Guilherme H. Travassos, Erika Nina Höhn, and Victor R. Basili. <i>A framework for software engineering experimental replications</i> . In Proceedings of the 13th IEEE International Conference on on Engineering of Complex Computer Systems, pages 203-212, Washington, DC, USA, 2008. IEEE Computer Society.	Irrelevante
60	2007	EI compendex	Lionel C. Briand. <i>Empirical evaluation in software engineering: role, strategy, and limitations</i> . In Proceedings of the international conference on Empirical software engineering issues: critical assessment and future directions, pages 21-21, Berlin, Heidelberg, 2007. Springer-Verlag.	Incompleto
61	2004	EI compendex	Hyunsook Do, Sebastian Elbaum, and Gregg Rothermel. <i>Infrastructure support for controlled experimentation with software testing and regression testing techniques</i> . In International Symposium on Empirical Software Engineering (ISESE'04), pages 60-70, 2004.	Irrelevante*
62	2009	EI compendex	Vitor Pires Lopes and Guilherme Horta Travassos. <i>Knowledge repository structure of an experimental software engineering environment</i> . In Proceedings of the XXIII Brazilian Symposium on Software Engineering, SBES '09, pages 32-42, Washington, DC, USA, 2009. IEEE Computer Society.	Irrelevante
63	2007	EI compendex	Walter F. Tichy and Frank Padberg. <i>Empirical methods in software engineering research</i> . In International Conference on Software Engineering Companion, pages 163-164, Los Alamitos, CA, USA, 2007. IEEE Computer Society.	Estudo repetido, 1ª fonte: IEEE
64	2007	EI compendex	Victor R. Basili, Marvin V. Zelkowitz, Dag I. Sjøberg, Philip Johnson, and Anthony J. Cowling. <i>Protocols in the use of empirical software engineering artifacts</i> . Empirical Softw. Engg., 12:107-119, February 2007.	Irrelevante
65	2009	EI compendex	Helen M. Edwards, Sharon McDonald, and S. Michelle Young. <i>The repertory grid technique: Its place in empirical software engineering research</i> . Inf. Softw. Technol., 51(4):785-798, April 2009.	Estudo repetido, 1ª fonte: Science Direct

ID	Ano	Fonte	Referência	Critério
66	2005	EI compendex	Amela Karahasanoviæ, Bente Anda, Erik Arisholm, Siw Elisabeth Hove, Magne Jørgensen, Dag I. Sjøberg, and Ray Welland. <i>Collecting feedback during software engineering experiments</i> . Empirical Softw. Engg., 10:113-147, April 2005.	Irrelevante*
67	2002	EI compendex	Barbara A. Kitchenham, Shari Lawrence Pfleeger, Lesley M. Pickard, Peter W. Jones, David C. Hoglin, Khaled El Emam, and Jarrett Rosenberg. <i>Preliminary guidelines for empirical research in software engineering</i> . IEEE Trans. Softw. Eng., 28:721-734, August 2002.	Estudo repetido, 1ª fonte: IEEE
68	2004	EI compendex	Andreas Jedlitschka and Marcus Ciolkowski. <i>Towards evidence in software engineering</i> . In Proceedings of the International Symposium on Empirical Software Engineering, pages 261-270, Washington, DC, USA, 2004. IEEE Computer Society.	Estudo repetido, 1ª fonte: IEEE
69	2007	EI compendex	Stefan Biffi and Dietmar Winkler. <i>Value-based empirical research plan evaluation</i> . In Proceedings of the First International Symposium on Empirical Software Engineering and Measurement, ESEM '07, pages 494-, Washington, DC, USA, 2007. IEEE Computer Society.	Irrelevante
70	2007	EI compendex	James Miller. <i>Creating real value in software engineering experiments</i> . In Proceedings of the international conference on Empirical software engineering issues: critical assessment and future directions, pages 38-38, Dagstuhl Castle, Germany, 2006. Springer-Verlag.	Estudo repetido, 1ª fonte: Scopus
71	2009	EI compendex	Gregory Gay, Tim Menzies, Bojan Cukic, and Burak Turhan. <i>How to build repeatable experiments</i> . In Proceedings of the 5th International Conference on Predictor Models in Software Engineering, PROMISE '09, pages 15:1-15:9, New York, NY, USA, 2009. ACM.	Estudo repetido, 1ª fonte: ACM
72	1996	EI compendex	Michael Angel. <i>Conducting experiments with experiment manager</i> . In Proceedings of the 28th conference on Winter simulation, WSC '96, pages 535-541, Washington, DC, USA, 1996. IEEE Computer Society.	Irrelevante
73	1997	EI compendex	Chris F. Kemerer and Sandra Slaughter. <i>Methodologies for performing empirical studies: Report from the international workshop on empirical studies of software maintenance</i> . Empirical Software Engineering, 2(2):109-118, 1997.	Irrelevante

ID	Ano	Fonte	Referência	Critério
74	2010	EI compendex	Roel Wieringa. <i>Design science methodology: principles and practice</i> . In Proceedings of the 32nd ACM/IEEE International Conference on Software Engineering - Volume 2, ICSE '10, pages 493-494, New York, NY, USA, 2010. ACM.	Estudo repetido, 1ª fonte: ACM
75	2005	EI compendex	Forrest Shull, Jeffrey Carver, Lorin Hochstein, and Victor Basili. <i>Empirical study design in the area of high performance computing (HPC)</i> . In International Symposium on Empirical Software Engineering (ISESE'05), pages 305-314, 2005.	Irrelevante*
76	2007	EI compendex	Carolyn B. Seaman. <i>Empirical paradigm: position paper</i> . In Proceedings of the international conference on Empirical software engineering issues: critical assessment and future directions, pages 23-23, Berlin, Heidelberg, 2007. Springer-Verlag.	Incompleto
77	2004	EI compendex	Dewayne E. Perry, Susan Elliott Sim, and Steve M. Easterbrook. <i>Case studies for software engineers</i> . In Proceedings of the 26th International Conference on Software Engineering, ICSE '04, pages 736-738, Washington, DC, USA, 2004. IEEE Computer Society.	Incompleto
78	2005	EI compendex	A.A. Naurnov and S. Gupta. <i>Some approaches and techniques for efficient control of experiments</i> . In Proceedings. The 9th Russian-Korean International Symposium on, pages 656-659, 2005.	Irrelevante
79	1999	EI compendex	Victor R. Basili, Forrest Shull, Filippo Lanubile, and IEEE Computer Society. <i>Building knowledge through families of experiments</i> . IEEE Transactions on Software Engineering, 25:456-473, 1999.	Estudo repetido, 1ª fonte: IEEE
80	2007	ESEM	Martin Host and Per Runeson. <i>Checklists for software engineering case study research</i> . In Proceedings of the First International Symposium on Empirical Software Engineering and Measurement, ESEM '07, pages 479-481, Washington, DC, USA, 2007. IEEE Computer Society.	Estudo repetido, 1ª fonte: IEEE
81	2008	EASE	Francisco Novillo, Félix Garcia, Elvira Rolón, Francisco Ruiz, and Mario Piattini. <i>Empirical webgen; a web-based environment for the automatic generation of surveys and experiments</i> . In 12th International Conference on Evaluation and Assessment in Software Engineering (EASE), University of Bari, Italy, Jun 2008.	Irrelevante

ID	Ano	Fonte	Referência	Critério
82	2008	ESEJ	Barbara Kitchenham, Hiyam Al-Khilidar, Muhammed Ali Babar, Mike Berry, Karl Cox, Jacky Keung, Felicia Kurniawati, Mark Staples, He Zhang, and Liming Zhu. <i>Evaluating guidelines for reporting empirical software engineering studies</i> . Empirical Softw. Engg. Journal, 13:97-121, February 2008.	Estudo repetido, 1ª fonte: Scopus
83	2009	ESEJ	Per Runeson and Martin Höst. <i>Guidelines for conducting and reporting case study research in software engineering</i> . Empirical Softw. Engg. Journal, 14:131-164, April 2009.	Estudo repetido, 1ª fonte: Ei compendex
84	2008	ESEJ	Tim Menzies. <i>Editorial, special issue, repeatable experiments in software engineering</i> . Empirical Softw. Engg. Journal, 13(5):469-471, October 2008.	Incompleto
85	2008	ESEJ	Yvonne Dittrich, Kari Rönkkö, Jeanette Eriksson, Christina Hansson, and Olle Lindeberg. <i>Cooperative method development. combining qualitative empirical research with method, technique and process improvement</i> . Empirical Software Engineering Journal, 13:231-260, 2008.	Estudo repetido, 1ª fonte: Ei compendex
86	2007	ESEJ	Victor R. Basili, Marvin V. Zelkowitz, Dag I. Sjøberg, Philip Johnson, and Anthony J. Cowling. <i>Protocols in the use of empirical software engineering artifacts</i> . Empirical Softw. Engg. Journal, 12:107-119, February 2007.	Estudo repetido, 1ª fonte: Ei compendex
87	2005	ESEJ	Hyunsook Do, Sebastian Elbaum, and Gregg Rothermel. <i>Supporting controlled experimentation with testing techniques: An infrastructure and its potential impact</i> . Empirical Softw. Engg., 10(4):405-435, October 2005.	Estudo repetido, 1ª fonte: Ei compendex
88	2005	ESEJ	Timothy C. Lethbridge, Susan Elliott Sim, and Janice Singer. <i>Studying software engineers: Data collection techniques for software field studies</i> . Empirical Softw. Engg. Journal, 10(3):311-341, July 2005.	Irrelevante*
89	2004	ESEJ	Forrest Shull, Manoel G. Mendonça, Victor Basili, Jeffrey Carver, José C. Maldonado, Sandra Fabbri, Guilherme Horta Travassos, and Maria Cristina Ferreira. <i>Knowledge-sharing issues in experimental software engineering</i> . Empirical Softw. Engg., 9:111-137, March 2004.	Estudo repetido, 1ª fonte: Ei compendex

ID	Ano	Fonte	Referência	Critério
90	2001	ESEJ	Susan Elliott Sim, Janice Singer, and Margaret-Anne Storey. <i>Beg, borrow, or steal: Using multidisciplinary approaches in empirical software engineering research</i> . Empirical Softw. Engg. Journal, 6(1):85-93, March 2001.	Incompleto
91	1999	ESEJ	R. Harrison, N. Badoo, E. Barry, S. Biff, A. Parra, B. Winter, and J. Wuest. <i>Directions and methodologies for empirical software engineering research</i> . Empirical Softw. Engg. Journal, 4(4):405-410, December 1999.	Irrelevante

APÊNDICE C

PROTOCOLO DO MAPEAMENTO SISTEMÁTICO

C.1 EQUIPE

- Emanuel Francisco Spósito Barreiros (Author)
 - Contact: efsb@cin.ufpe.br
 - Affiliation: Informatics Center - Federal University of Pernambuco
- Sérgio Castelo Branco Soares (Revisor e Autor)
 - Contact: scbs@cin.ufpe.br
 - Affiliation: Informatics Center - Federal University of Pernambuco
- Adauto Trigueiro de Almeida Filho (Revisor e Autor)
 - Contact: ataf@cin.ufpe.br
 - Affiliation: Informatics Center - Federal University of Pernambuco
- Juliana Gonçalves Saraiva (Revisor)
 - Contact: jags2@cin.ufpe.br
 - Affiliation: Informatics Center - Federal University of Pernambuco

C.2 INTRODUÇÃO

As Revisões Sistemáticas (RS) são métodos de pesquisa eficazes para reunião de conhecimentos sobre uma determinada área ou tópico de pesquisa, e, pelo menos dois tipos podem ser encontradas na literatura [47], uma delas é a revisão sistemática da literatura convencional que tem como foco agregar resultados sobre a eficácia de um determinado tratamento, intervenção ou tecnologia [47], ela é um dos principais métodos de pesquisa da Engenharia de Software Baseada em Evidências (ESBE) [41] e, diferentemente de um revisão comum, ela tem que explicitamente ter uma estratégia de pesquisa e critérios de aceitação, que permitam que as evidências pertinentes sejam consideradas de forma

sistemática e transparente. O outro tipo de RS que pode ser encontrado é o Mapeamento Sistemático (MS), ou como também é conhecido Estudo de Escopo (EE), algumas definições sobre MS são apresentadas abaixo:

- MS é um método de pesquisa que prover uma visão geral de uma área de pesquisa e nos permite identificar a quantidade o o tipo de pesquisa e resultados disponíveis dentro dela [44].
- EE pode ter como objetivo mapear rapidamente os principais conceitos que sustentam uma área de pesquisa, as principais fontes e os tipos de evidências disponíveis, ele pode ser realizado de forma autônoma, especialmente quando uma área é complexa ou não tenha sido revisada exaustivamente [58].
- MS é uma revisão ampla dos estudos primários de uma determinada área que visa identificar as evidências disponíveis no tópico investigado [41].

Em resumo, o MS tem como objetivo mapear uma determinada área de pesquisa de forma que se descubra o que está sendo estudado e também lacunas, ou seja, onde não há estudos sendo conduzidos. Um MS deve seguir etapas bem definidas, que são elaboradas previamente em um protocolo, isso é feito para manter a clareza da execução do MS e para tornar possível a replicação do estudo por outros pesquisadores.

Askey e O'Malley [48], apresentam alguns motivos para a execução de um MS:

- Para examinar a extensão, alcance e natureza dos fenômenos de investigação. Este tipo de revisão é útil para mapeamento de áreas de estudo, onde é difícil visualizar a gama de materiais que possam estar disponíveis;
- Para determinar a necessidade de realizar uma revisão sistemática convencional. Nestes casos, um mapeamento preliminar da literatura pode ser realizado para identificar se é viável ou relevante executar uma revisão sistemática convencional;
- Para resumir e divulgar resultados de pesquisas. Esse tipo de estudo de escopo pode descrever em detalhes os resultados e a abrangência da pesquisa em determinadas áreas de estudo, proporcionando assim um mecanismo de síntese e divulgação dos resultados da investigação;
- Para identificar as lacunas de pesquisa na literatura existente. Este tipo de estudo de escopo elabora conclusões, a partir da literatura existente, sobre o estado global da área de pesquisa investigada. Especificamente concebido para identificar lacunas

na base de evidências onde não há pesquisas sendo conduzidas, o estudo de escopo pode também resumir e disseminar resultados de pesquisas, bem como, identificar a relevância de uma revisão sistemática convencional em uma área de estudo específica. No entanto é importante notar que identificar lacunas na literatura através do estudo de escopo não é necessariamente identificar lacunas na pesquisa em si, ou seja, identificar se um pesquisa tem uma qualidade pobre, uma vez que a avaliação da qualidade não faz parte da missão de um estudo de escopo.

Este documento apresenta o protocolo deste MS, cujo objetivo é identificar os principais mecanismos usados para guiar estudos empíricos em engenharia de software, possibilitando dessa forma reunir tal conhecimento para que a comunidade científica conheça formas de guiar seus estudos empíricos, e que no futuro haja vários estudos metodologicamente semelhantes, e assim, sejam passíveis de comparação para contribuir para a construção de um corpo de conhecimento na área.

C.3 ESCOPO E QUESTÕES DE PESQUISA

Com o objetivo de delinear o escopo da pesquisa e de identificar os elementos que virá a fazer parte das questões de pesquisa, foi utilizado a estrutura PICOC [41].

- **População (Population):** Pesquisas que conduzem estudos empíricos em engenharia de software.
- **Intervenção (Intervention):** uso de mecanismos para guiar estudos empíricos.
- **Comparação (Comparison):** esta parte da estrutura não foi utilizada.
- **Resultado (Outcomes):** Um documento que mapeia os mecanismos para guiar estudos empíricos em engenharia de software.
- **Contexto (Context):** indústria e academia.

O item Comparação (Comparison) não foi utilizado, uma vez que o estudo não realiza comparações entre os mecanismos para guiar estudos empíricos. O item da estrutura denominado contexto (Context) é normalmente é utilizado para definir o contexto que ocorre a comparação. No entanto, aqui será utilizado para definir o contexto admitido para os mecanismos encontrados.

Após a definição da estrutura PICOC as perguntas que guiaram o estudo foram definidas.

- **Questão 1:** Quais são os mecanismos existentes para guiar estudos empíricos em engenharia de software?
- **Questão 2:** Os mecanismos encontrados são utilizados para guiar que tipo de estudo empírico?
- **Questão 3:** Quais elementos, fases, etapas compõem os mecanismos utilizados para guiar estudos empíricos?

C.4 ESTRATÉGIA E FONTES DE BUSCA

A construção da *string* de busca utilizada nas bibliotecas digitais selecionadas seguiu a seguinte estratégia:

- Derivar a partir das questões de pesquisa as principais palavras-chaves a partir da identificação da população, intervenção, comparação (quando for o caso), resultados e contexto;
- Procurar por palavras chaves em artigos relevantes já consultados em uma revisão informal;
- Identificar sinônimos e termos alternativos as palavras chaves;
- Consultar especialistas na área;
- Usar o conector booleano OR para incorporar palavras alternativas e sinônimas.
- usar o conector booleano AND para ligar palavras chaves
- Verificar a string de busca construída realizando buscas piloto e comparando os resultados obtidos com uma lista de estudos primários já conhecidos;

Como resultado da estratégia supracitada foi obtido à seguinte string de busca:

(framework OR method OR methodology OR guideline OR process OR report OR reporting OR support OR supporting OR checklist OR tutorial) AND (empirical OR experiment OR experimentation OR experimental OR survey OR “case study” OR “action research” OR ethnography OR “controlled experiment” OR quasi-experiment OR “quasi experiment”) AND (“software engineering”)

A procura por estudos primários na literatura incluiu buscas automatizadas através de engenhos de busca das bibliotecas digitais, onde foi utilizada a *string* de busca formulada e buscas manuais, com o objetivo de ampliar a cobertura da pesquisa e dar mais segurança ao pesquisador.

As bibliotecas digitais utilizadas na busca automatizada foram:

- IEEE Computer Society Digital Library;
- ACM;
- Science Direct;
- Scopus;
- Ei Compendex.

As conferências utilizadas na busca manual foram:

- Empirical Software Engineering and Measurement (ESEM), no período de 2007 a 2010;
- Evaluation and Assessment in Software Engineering (EASE), no período de 2006 a 2010.

E por fim, o periódico utilizado na busca manual foi:

- Empirical Software Engineering Journal (ESEJ), no período de 1997 a 2010.

C.5 SELEÇÃO DE ESTUDOS

C.5.1 Processo Modelo de Seleção

A seleção dos estudos primários que fazem parte deste mapeamento seguiram um processo modelo que pode ser visto na Figura C.1. O processo parte sempre de um conjunto de estudos que passam por revisores (representado pela letra R) e em seguida os seus resultados agregados em um tabela de consensos e conflitos, ou seja, onde os revisores concordam ou discordam sobre a inserção de um estudo no mapeamento. Logo após a geração da tabela há uma reunião para resolução dos possíveis conflitos e geração do conjunto final de estudos.

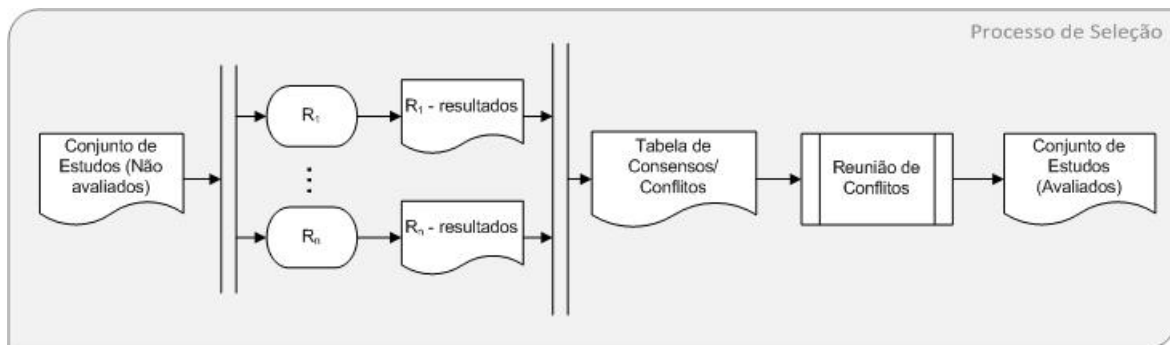


Figura C.1 Processo modelo usado para seleção de estudos primários

C.5.2 Critérios de Inclusão/Exclusão

A inclusão de um trabalho no mapeamento se dá pela sua relevância em relação às questões de investigação. Os critérios de inclusão adotados pelo estudo foram:

- Estudos que definem um mecanismo para guiar estudos empíricos;
- Estudos que estão relacionados a engenharia de software.

Os critérios de exclusão adotados pelo estudo foram:

- Estudos irrelevantes para a pesquisa;
- Estudos que são repetidos, se um estudo estiver disponível em diferentes fontes, a primeira pesquisa que será considerada;
- Estudos Duplicados, no caso de haver dois ou mais estudos semelhantes apenas o mais completo será considerado, exceto em situações onde há informações adicionais.
- Estudos incompletos como resumos ou resumos expandidos e apresentações (Power Point).

C.5.3 Etapas de Seleção dos Estudos

A seleção dos estudos primário ocorreram em duas etapas e as mesmas seguiram o processo de seleção proposto na Seção C.5.1. A primeira etapa, ilustrada pela Figura C.2, segue os seguintes passos:

- Obter um primeiro conjunto de estudos, não avaliados, através dos engenhos de buscas das bibliotecas digitais e da busca manual.
- Os revisores EB e AA selecionam potenciais estudos primários através da leitura do título, resumo e palavras-chave. Neste passo, estudos completamente irrelevantes são excluídos do estudo.
- Os resultado do passo anterior é reunido em uma tabela onde os consensos e os conflitos são expostos.
- Há uma reunião para resolução de conflitos, a reunião ainda se concentra em aspectos superficiais dos estudos, pois ainda não foram lidos na íntegra.
- Por fim, um conjunto de potenciais estudos é gerado.

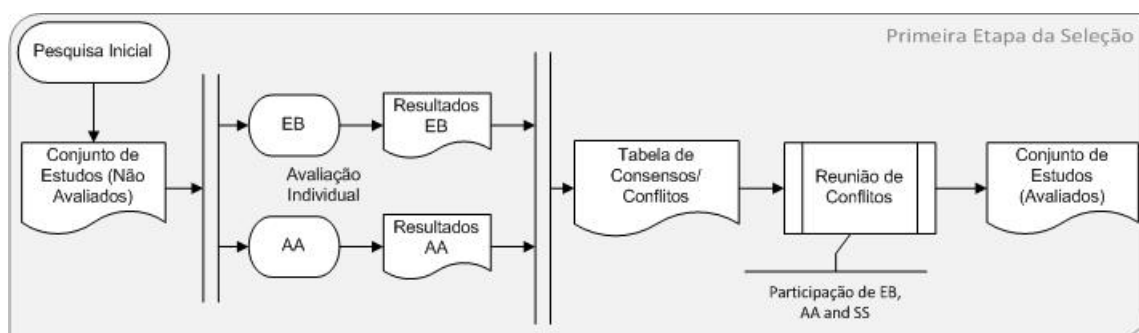


Figura C.2 Primeira Etapa da Seleção

Após o fim da primeira etapa, um conjunto de potenciais estudos é gerado e a segunda etapa da seleção, ilustrada pela Figura C.3, é iniciada. Ela é descrita a seguir:

- Com o conjunto de potenciais estudos primários, é formada duas duplas, uma composta por AA e JS e outra por EB e JS. Cada uma das duplas é responsável por ler na íntegra metade dos potenciais estudos selecionados na etapa anterior.
- Após a leitura os resultados sobre a decisão se o estudo está incluído ou excluído do mapeamento é reunido novamente em uma tabela de consensos e conflitos.
- Uma última reunião é realizada com todos os participantes (AA, EB, JS e SS) para resolução dos possíveis conflitos.
- Por fim o conjunto com os estudos primários selecionados é gerado.

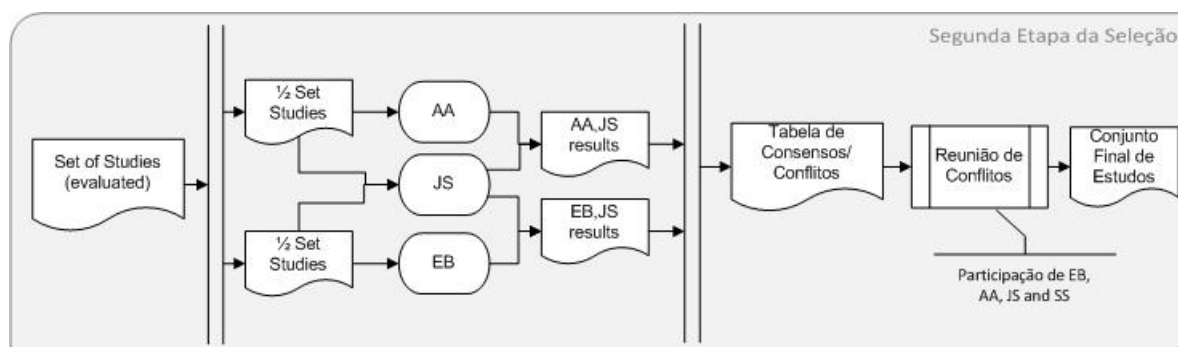


Figura C.3 Segunda Etapa da Seleção

C.6 EXTRAÇÃO DOS DADOS

To help the extraction of information from the selected and excluded studies, three forms were used. They are shown in the following sections.

Com a finalidade de extrair informações dos estudos primários selecionados, três formulários foram utilizados. Eles são apresentados nas seções seguintes.

C.6.1 Formulário A

Este formulário tem o objetivo de coletar informações gerais dos estudos primários. As seguintes informações são coletadas:

- **ID:** um identificador do estudo. Útil para referenciar o estudo em várias partes do mapeamento;
- **Fonte:** a fonte onde foi encontrado o estudo;
- **Ano:** o ano de publicação;
- **Título:** o título do estudo;
- **Autor:** a lista de autores do estudo;
- **País:** a lista dos países das instituições de origem dos autores.

C.6.2 Formulário B

Os estudos excluídos também precisam ser documentados. Uma das informações mais importantes a ser documentada neste caso é o motivo pelo qual o estudo foi excluído. As informações extraídas são as seguintes:

- **ID:** um identificador do estudo. Útil para referenciar o estudo em várias partes do mapeamento;
- **Fonte:** a fonte onde foi encontrado o estudo;
- **Ano:** o ano de publicação;
- **Título:** o título do estudo;
- **Autor:** a lista de autores do estudo;
- **Critério usado para exclusão:** O critério usado pelos revisores para excluir o estudo do mapeamento.

C.6.3 Formulário C

O formulário C é usado para extrair informações relevantes, a partir dos estudos primários selecionados, para responder as perguntas de pesquisa. As informações coletadas são:

- **Data da Avaliação:** quando o artigo foi avaliado;
- **Revisor:** pesquisador que extraiu as informações e preencheu os formulários;
- **Pergunta de Pesquisa 1:** quais são os mecanismos existentes para guiar estudos empíricos em engenharia de software?
- **Pergunta de Pesquisa 2:** os mecanismos encontrados são utilizados para guiar que tipo de estudo empírico?
- **Pergunta de Pesquisa 3:** quais elementos, fases, etapas compõem os mecanismos utilizados para guiar estudos empíricos?
- **Nota:** algo relevante que o revisor deseje documentar.

C.7 ESTRATÉGIA DE SÍNTESE DOS DADOS

A contribuição idealizada por este estudo é gerada a partir da síntese dos dados obtidos através do mapeamento sistemático realizado. O procedimento para a síntese dos dados se inicia com a marcação das passagens retiradas dos estudos selecionados (dados qualitativos), a partir dos formulários apresentados na Seção C.6, que fornecem algum tipo de informação relevante para responder as perguntas de pesquisa. Cada um dos trechos

retirados dos estudos primário é associado a uma código que identifica o tipo de estudo empírico e ou o tipo de mecanismo que o estudo está associado. Os códigos devem seguir um padrão para posteriormente serem agrupados. Por exemplo, nesta pesquisa foram extraídos passagens que indicam o tipo de mecanismo desenvolvido, qual tipo de estudo o mecanismo guia e a forma que está estruturado o guia. Os códigos deste estudo seguem a seguinte formatação:

- **ID_TM**;
- **ID_TEG**.

Onde:

- **ID** = **I**dentificador do estudo primário;
- **TM** = **T**ipo do **M**ecanismo encontrado para guiar os estudos empíricos em engenharia de software;
- **TEG** = **T**ipo do **E**studo **G**uiado pelo mecanismo desenvolvido.

Assim, pode-se agrupar os trechos das pesquisas analisadas em grupos de acordo com os códigos utilizados. O primeiro padrão de código identificou os trechos dos estudos que indicava o tipo de mecanismo desenvolvido, o segundo código identificou os trechos que apontavam o tipo de estudo guiado pelo mecanismo e por fim o terceiro código foi utilizado para identificar os trechos dos estudos que descreviam a estrutura do mecanismo desenvolvido.

Desta forma, seguindo o procedimento descrito e utilizando do auxílio de planilhas eletrônicas, foram mapeadas as evidências dos estudos selecionados e sintetizados os resultados apresentados no próximo capítulo. Todo processo foi realizado de forma que as informações possam ser rastreadas, possibilitando identificar os trabalhos que proveram os dados.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Manoel Mendonça, Daniela Cruzes, Josemeire Dias, and Maria Cristina Ferreira de Oliveira. Using observational pilot studies to test and improve lab packages. In *Proceedings of the 2006 ACM/IEEE international symposium on Empirical software engineering*, ISESE '06, pages 48–57, New York, NY, USA, 2006. ACM. [\(document\)](#), [4.2.3](#), [4.10](#)
- [2] Teade Punter, Marcus Ciolkowski, Bernd Freimut, and Isabel John. Conducting on-line surveys in software engineering. In *Proceedings of the 2003 International Symposium on Empirical Software Engineering*, ISESE '03, page 80, Washington, DC, USA, 2003. IEEE Computer Society. [\(document\)](#), [4.11](#)
- [3] V R Basili, R W Selby, and D H Hutchens. Experimentation in software engineering. *IEEE Trans. Softw. Eng.*, 12(7):733–743, 1986. [\(document\)](#), [1.1](#), [2.2](#), [4.12](#)
- [4] Miguel Goulao and Fernando Brito e. Abreu. Modeling the experimental software engineering process. In *Proceedings of the 6th International Conference on Quality of Information and Communications Technology*, pages 77–90, Washington, DC, USA, 2007. IEEE Computer Society. [\(document\)](#), [4.13](#)
- [5] Victor R. Basili and Richard W. Selby. Paradigms for experimentation and empirical studies in software engineering. *Reliability Engineering and System Safety*, 32(1-2):171–191, 1991. [\(document\)](#), [4.14](#)
- [6] null Shaochun Xu and V. Rajlich. Dialog-based protocol: an empirical research method for cognitive activities in software engineering. *Empirical Software Engineering, International Symposium on*, 0:10, 2005. [\(document\)](#), [4.15](#)
- [7] Barbara Kitchenham, Lesley Pickard, and Shari Lawrence Pfleeger. Case studies for method and tool evaluation. *IEEE Softw.*, 12:52–62, July 1995. [1.1](#), [2.2](#)
- [8] Koji Torii, Ken-ichi Matsumoto, Kumiyo Nakakoji, Yoshihiro Takada, Shingo Takada, and Kazuyuki Shima. Ginger2: An environment for computer-aided empirical software engineering. *IEEE Trans. Softw. Eng.*, 25:474–492, July 1999. [1.1](#), [2.2](#)

- [9] Claes Wohlin, Per Runeson, Martin Höst, Magnus C. Ohlsson, Björn Regnell, and Anders Wesslén. *Experimentation in software engineering: an introduction*. Kluwer Academic Publishers, Norwell, MA, USA, 2000. [1.1](#), [2.2](#), [4.2.3](#)
- [10] N. Juristo and A. Moreno. *Basics of Software Engineering Experimentation*. Kluwer Academic Publishers, Norwell, MA, USA, 2001. [1.1](#), [2.2](#)
- [11] Barbara A. Kitchenham, Shari Lawrence Pfleeger, Lesley M. Pickard, Peter W. Jones, David C. Hoaglin, Khaled El Emam, and Jarrett Rosenberg. Preliminary guidelines for empirical research in software engineering. *IEEE Trans. Softw. Eng.*, 28:721–734, August 2002. [1.1](#), [2.2](#)
- [12] Barbara A. Kitchenham, Tore Dyba, and Magne Jorgensen. Evidence-based software engineering. In *ICSE '04: Proceedings of the 26th International Conference on Software Engineering*, pages 273–281, Washington, DC, USA, 2004. IEEE Computer Society. [1.1](#), [2.2](#), [2.4](#), [3.2.1](#)
- [13] S. Easterbrook, J. Singer, M.A. Storey, and D. Damian. *Selecting Empirical Methods for Software Engineering Research*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2007. [1.1](#), [2.2](#), [2.3](#)
- [14] Per Runeson and Martin Höst. Guidelines for conducting and reporting case study research in software engineering. *Empirical Softw. Eng.*, 14:131–164, April 2009. [1.1](#), [2.2](#)
- [15] Shari Lawrence Pfleeger. Albert einstein and empirical software engineering. *Computer*, 32:32–38, October 1999. [1.1](#)
- [16] Victor R. Basili, Forrest Shull, Filippo Lanubile, and Ieee Computer Society. Building knowledge through families of experiments. *IEEE Transactions on Software Engineering*, 25:456–473, 1999. [1.1](#)
- [17] Victor R. Basili. The experimental paradigm in software engineering. In *Proceedings of the International Workshop on Experimental Software Engineering Issues: Critical Assessment and Future Directions*, pages 3–12, London, UK, 1993. Springer-Verlag. [1.1](#), [2.2](#)
- [18] Colin Robson. *Real World Research: A Resource for Social Scientists and Practitioners- Researchers*. Wiley-Blackwell, Malden, MA, USA, 2002. [1.1](#), [2.2](#)

- [19] E.A.G.G. AMARAL. Empacotamento de experimentos em engenharia de software. Master's thesis, University of Rio de Janeiro, 2003. [1.1](#), [2.2](#)
- [20] Susan Elliott Sim, Steve Easterbrook, and Richard C. Holt. Using benchmarking to advance research: a challenge to software engineering. In *ICSE '03: Proceedings of the 25th International Conference on Software Engineering*, pages 74–83, Washington, DC, USA, 2003. IEEE Computer Society. [1.1](#)
- [21] Walter F. Tichy. Should computer scientists experiment more? *Computer*, 31(5):32–40, 1998. [1.1](#), [2.2](#)
- [22] Susan Elliott Sim, Richard C. Holt, and Steve Easterbrook. On using a benchmark to evaluate c++ extractors. In *Proceedings of the 10th International Workshop on Program Comprehension, IWPC '02*, page 114, Washington, DC, USA, 2002. IEEE Computer Society. [1.1](#)
- [23] Marcelo Moura. Um benchmarking framework para avaliação da manutenibilidade de software orientado a aspectos. Master's thesis, University of Pernambuco, Recife, PE, Brasil, 2008. [1.1](#), [1.2.1](#), [2.2](#), [3.2](#)
- [24] Fritz Bauer. Software engineering. In Peter Naur and Brian Randell, editors, *A Report on a Conference Sponsored by the NATO Science Committee*, 1969. [2.1](#)
- [25] IEEE. Ieee glossary of software engineering terminology, ieee standard 610.12. Technical report, IEEE, 1990. [2.1](#)
- [26] Ian Sommerville. *Software Engineering*. Addison-Wesley Publishing Company, USA, 2007. [2.1](#), [4.2.1](#)
- [27] Dag I. K. Sjoberg, Jo E. Hannay, Ove Hansen, Vigdis By Kampenes, Amela Karahasanovic, Nils-Kristian Liborg, and Anette C. Rekdal. A survey of controlled experiments in software engineering. *IEEE Trans. Softw. Eng.*, 31:733–753, September 2005. [2.2](#)
- [28] John W. Daly, Khaled E. Emam, and James Miller. Multi-method research in software engineering. In *IEEE Workshop on Empirical Studies Software Maintenance*, Bari, Italy, 1997. [2.2](#)
- [29] Victor R. Basili. The role of experimentation in software engineering: past, current, and future. In *ICSE '96: Proceedings of the 18th international conference on Software*

- engineering*, pages 442–449, Washington, DC, USA, 1996. IEEE Computer Society. 2.2
- [30] Dewayne E. Perry, Adam A. Porter, and Lawrence G. Votta. Empirical studies of software engineering: a roadmap. In *ICSE '00: Proceedings of the Conference on The Future of Software Engineering*, pages 345–355, New York, NY, USA, 2000. ACM. 2.2
- [31] Walter F. Tichy, Paul Lukowicz, Lutz Prechelt, and Ernst A. Heinz. Experimental evaluation in computer science: a quantitative study. *J. Syst. Softw.*, 28(1):9–18, 1995. 2.2
- [32] Marvin V. Zelkowitz and Dolores R. Wallace. Experimental models for validating technology. *Computer*, 31(5):23–31, 1998. 2.2
- [33] Samuel T. Redwine, Jr. and William E. Riddle. Software technology maturation. In *ICSE '85: Proceedings of the 8th international conference on Software engineering*, pages 189–200, Los Alamitos, CA, USA, 1985. IEEE Computer Society Press. 2.2
- [34] Andreas Jedlitschka, Marcus Ciolkowski, Christian Denger, Bernd Freimut, and Andreas Schlichting. Relevant information sources for successful technology transfer: A survey using inspections as an example. In *ESEM '07: Proceedings of the First International Symposium on Empirical Software Engineering and Measurement*, pages 31–40, Washington, DC, USA, 2007. IEEE Computer Society. 2.2
- [35] Bent Flyvbjerg. Five misunderstandings about case-study research. *Qualitative Inquiry*, 12(2):219–245, 2006. 2.3
- [36] Robert K. Yin. *Case Study Research: Design and Methods*. Applied social research methods series. Sage Publications, Beverly Hills, CA, 1984. 2.3
- [37] Hugh Robinson, Judith Segal, and Helen Sharp. Ethnographically-informed empirical studies of software practice. *Inf. Softw. Technol.*, 49:540–551, June 2007. 2.3
- [38] Robert M. Davison, Maris G. Martinsons, and Ned Kock. Principles of canonical action research. *Information Systems Journal*, 14:65–86, 2004. 2.3
- [39] Tore Dyba, Torgeir Dingsoyr, and Geir K. Hanssen. Applying systematic reviews to diverse study types: An experience report. In *Proceedings of the First International Symposium on Empirical Software Engineering and Measurement*, ESEM '07, pages 225–234, Washington, DC, USA, 2007. IEEE Computer Society. 2.4, 2.4

- [40] Graham Capper Briony J Oates. Using systematic reviews and evidence-based software engineering with masters students. In *EASE '09: 13th International Conference on Evaluation and Assessment in Software Engineering*. BCS, 2009. 2.4, 2.4
- [41] B. A. Kitchenham and S. Charters. Guidelines for performing systematic literature reviews in software engineering. Technical Report EBSE 2007-001, Keele University and Durham University Joint Report, July 2007. 2.4, 2.4, 3.1, 3.2.1, 3.2.2, 4.1, 5.1, C.2, C.3
- [42] Biolchini J. Travassos, G. Revisões sistemáticas aplicadas a engenharia de software. In *XXI SBES - Brazilian Symposium on Software Engineering*, João Pessoa, PB, Brasil, 2007. 2.4, 2.4
- [43] Tore Dyba, Barbara A. Kitchenham, and Magne Jorgensen. Evidence-based software engineering for practitioners. *IEEE Software*, 22:58–65, January 2005. 2.4
- [44] Kai Petersen, Robert Feldt, Shahid Mujtaba, and Michael Mattsson. Systematic mapping studies in software engineering. *12th International Conference on Evaluation and Assessment in Software Engineering*, pages 71–80, 2008. 2.4, C.2
- [45] Jo E. Hannay, Dag I.K. Sjøberg, and Tore Dyba. A systematic review of theory use in software engineering experiments. *IEEE Transactions on Software Engineering*, 33:87–107, 2007. 2.4
- [46] Tore Dybå, Vigdis By Kampenes, and Dag I.K. Sjøberg. A systematic review of statistical power in software engineering experiments. *Information and Software Technology*, 48(8):745 – 755, 2006. 2.4
- [47] M. Petticrew and H. Roberts. *Systematic Reviews in the Social Sciences: A Practical Guide*. Wiley-Blackwell, 2005. 2.4, 3.2.2, C.2
- [48] Hilary Arksey and Lisa O'Malley. Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology*, 8(1):19–32, February 2005. 2.4, C.2
- [49] Eva Maria Lakatos and Maria de Andrade Marconi. *Metodologia Científica*. Atlas, 2007. 3.1, 4.2.1
- [50] Harris M. Cooper. Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge, Technology & Policy*, 1(1):104–126, 1988. 3.1.1, 3.1.1

- [51] Barbara Kitchenham, Emilia Mendes, and Guilherme H. Travassos. A systematic review of cross- vs. within-company cost estimation studies. In *Proceedings of EASE 2006*. British Informatics Society Ltd, pages 10–12, 2006. [3.2.3](#)
- [52] Ralph E. Johnson and Brian Foote. Designing Reusable Classes. *Journal of Object-Oriented Programming*, 1(2):22–35, 1988. [4.2.1](#)
- [53] Darren Govoni. *Java application frameworks*. John Wiley & Sons, Inc., New York, NY, USA, 1999. [4.2.1](#)
- [54] Thomas S. Kuhn. *The Structure of Scientific Revolutions*. University Of Chicago Press, 1962. [4.2.1](#)
- [55] Martin Host and Per Runeson. Checklists for software engineering case study research. In *Proceedings of the First International Symposium on Empirical Software Engineering and Measurement*, ESEM '07, pages 479–481, Washington, DC, USA, 2007. IEEE Computer Society. [4.2.1](#), [4.2.2](#)
- [56] Mario Bunge. *Epistemologia: curso de atualização*. T. A. Queiroz/Edusp, 1980. [4.2.1](#)
- [57] Bernd Freimut, Teade Punter, Stefan Biffli, and Marcus Ciolkowski. State-of-the-art in empirical studies authors:. Technical Report VISEK/007/E, Fraunhofer IESE, march 2002. [4.2.3](#)
- [58] N Mays, E Roberts, and Jennie Popay. *Synthesising research evidence.*, pages 188–220. Routledge, London, UK, 2001. [C.2](#)

Este volume foi tipografado em L^AT_EX na classe UFPETthesis (www.cin.ufpe.br/~paguso/ufpethesis).
Para detalhes sobre este documento, clique [aqui](#).