



**Universidade de Brasília**

Instituto de Ciências Exatas  
Departamento de Ciência da Computação

**Ferramentas de Apoio a Experimentos: Um  
Mapeamento Sistemático**

Ricardo de Lima

Dissertação apresentada como requisito parcial para  
conclusão do Mestrado em Informática

Orientador  
Prof. Dr. Vander Ramos Alves

Brasília  
2018

Ficha catalográfica elaborada automaticamente,  
com os dados fornecidos pelo(a) autor(a)

LL732f Lima, Ricardo de  
Ferramentas de Apoio a Experimentos: Um Mapeamento  
Sistemático / Ricardo de Lima; orientador Vander Ramos  
Alves. -- Brasília, 2018.  
87 p.

Dissertação (Mestrado - Mestrado em Informática) --  
Universidade de Brasília, 2018.

1. experimentos controlados. 2. mapeamento sistemático.  
3. ferramentas de apoio. 4. engenharia de software. I.  
Alves, Vander Ramos, orient. II. Título.



# Dedicatória

Ao meu irmão, que sempre adiantado, passa pelas barreiras da vida antes de mim, e me apoia quando chega a minha vez.

A todos os nerds, que, como eu, aproveitam a vida do nosso jeito louco.

*"Sonho que se sonha só*

*É só um sonho que se sonha só*

*Mas sonho que se sonha junto é realidade"*

Raul Seixas

# Agradecimentos

Agradeço em primeiro lugar à Deus, por ter me dado forças para completar todas as tarefas que inicio. Em seguida, agradeço especialmente ao meu orientador, o professor Vander Alves, por ter me apoiado em todos esses anos até concluir o mestrado. Sem ele essa tarefa seria impossível.

Agradeço à minha família, por todas as orações que fizeram por mim, e que, mesmo distantes, me faziam sentir acolhido. Agradeço aos meus superiores em todos os lugares que trabalhei desde o início desse mestrado, pois a compreensão deles quanto ao meu cansaço no expediente e horários especiais que precisei me ausentar, foram essenciais para chegar até aqui.

Por fim, agradeço aos meus amigos, colegas de trabalho e todas as pessoas que estiveram envolvidas comigo durante essa trajetória.

# Resumo

Pesquisadores que desejam obter maior rigor em seus estudos e apresentar maior confiabilidade a seus resultados podem utilizar diversos métodos de estudos empíricos, dentre eles os experimentos controlados. Para isso existem guias que definem fases para um projeto de experimentação, e ferramentas são desenvolvidas para apoiar na realização das atividades. Nesse contexto, este trabalho tem como objetivo analisar estudos que apresentam ferramentas de apoio a experimentação de projetos em engenharia de software, e obter uma relação dessas ferramentas indicando quais técnicas, abstrações de linguagem, automação e garantias são oferecidos por elas em cada fase de um projeto de experimentos. O resultado será sintetizado para identificar os pontos fortes e fracos e quais os problemas mais comumente identificados entre elas. Em relação ao método, será realizado um mapeamento sistemático baseado em questões de pesquisa e um protocolo de pesquisa para extrair informações dos estudos primários. Em termos de resultados esperados, obtemos um guia para ajudar o pesquisador em engenharia de software a escolher uma ferramenta ao realizar experimentos.

**Palavras-chave:** experimentos controlados, mapeamento sistemático, ferramentas de apoio, engenharia de software

# Abstract

Researchers who wish to be more rigorous in their studies and present greater reliability to their results can use several methods of empirical studies, among them controlled experiments. There are guides that define phases for a project of experimentation, and tools are developed to support the accomplishment of the activities. In this context, this work aims to analyze studies that present tools to support the experimentation of projects in software engineering, and to obtain a list of these tools indicating which techniques, language abstractions, automation and guarantees are offered by them in each phase of a design of experiments. The result will be synthesized to identify the strengths and weaknesses and which are the most commonly identified problems among them. Regarding method, a systematic mapping will be performed based on research questions and a research protocol to extract information from the primary studies. In terms of expected results, we obtain a guide to help software engineering researcher choose a tool when conducting experiments.

**Keywords:** controlled experiments, systematic mapping, support tools, Software engineering

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Definição do Problema . . . . .	2
1.2	Solução Proposta . . . . .	2
1.3	Resumo das Contribuições . . . . .	3
1.4	Organização do Trabalho . . . . .	3
<b>2</b>	<b>Referencial Teórico</b>	<b>5</b>
2.1	Experimentos em Engenharia de Software . . . . .	5
2.1.1	Fase de Escopo . . . . .	7
2.1.2	Fase de Planejamento . . . . .	7
2.1.3	Fase de Execução . . . . .	9
2.1.4	Fase de Análise e Interpretação . . . . .	10
2.1.5	Fase de Apresentação e Empacotamento . . . . .	10
2.2	Processo de Revisão da Literatura . . . . .	11
2.2.1	Mapeamento Sistemático . . . . .	11
<b>3</b>	<b>Método de Pesquisa</b>	<b>13</b>
3.1	Questões de Pesquisa . . . . .	15
3.2	Estratégia de busca . . . . .	16
3.3	Critério de Seleção . . . . .	17
3.4	Processo de triagem . . . . .	18
3.5	Extração de Dados . . . . .	20
3.6	Análise e Síntese dos Dados . . . . .	20
<b>4</b>	<b>Resultados e Análise</b>	<b>22</b>
4.1	Resultados do Protocolo de Pesquisa . . . . .	22
4.2	Estudos Primários Seleccionados e Ferramentas Identificadas . . . . .	24

4.3	Resultados Sobre as Questões de Pesquisa . . . . .	28
4.3.1	Questão de Pesquisa 1 - Quais são as ferramentas de apoio à experimentação em Engenharia de Software e as funcionalidades de experimentação suportadas pelas ferramentas? . . . . .	28
4.3.2	Questão de Pesquisa 2 - Quais são as fases de um projeto de experimentação cobertas pela ferramenta? . . . . .	30
4.3.3	Questão de Pesquisa 2.1 - Qual o nível de abstração da linguagem de especificação da ferramenta? . . . . .	31
4.3.4	Questão de Pesquisa 2.2 - Qual é o nível de automação suportado pela ferramenta? . . . . .	33
4.3.5	Questão de Pesquisa 2.3 - Qual é a confiabilidade fornecida pela ferramenta? . . . . .	34
4.3.6	Questão de Pesquisa 3 - Que tipo de experimento é apoiado pela ferramenta? . . . . .	35
4.4	Análise Geral das Ferramentas . . . . .	36
4.4.1	Ferramentas de Apoio a Experimentos Não Baseados em Tecnologia . . . . .	37
4.4.2	Ferramentas de Apoio a Experimentos Baseados em Tecnologia . . . . .	40
4.4.3	Discussão e Implicações . . . . .	41
4.5	Guia Preliminar para Escolha de Ferramenta . . . . .	42
4.5.1	Identificação das Necessidades do Experimento e Suporte Oferecido . . . . .	43
4.5.2	Vantagens e Desvantagens das Ferramentas Identificadas . . . . .	46
<b>5</b>	<b>Conclusão, Trabalhos Relacionados, Ameaças a Validade e Trabalhos Futuros</b>	
	<b>Futuros</b>	<b>51</b>
5.1	Conclusão . . . . .	51
5.2	Trabalhos Relacionados . . . . .	53
5.3	Ameaças à Validade do Estudo . . . . .	55
5.4	Trabalhos Futuros . . . . .	56
	<b>Referências</b>	<b>57</b>
	<b>Anexo</b>	<b>59</b>
<b>I</b>	<b>Formulário de Extração de Dados</b>	<b>60</b>
<b>II</b>	<b>Formulários Preenchidos pela Extração de Dados</b>	<b>62</b>

# Lista de Figuras

2.1 Fases de um projeto de experimentos . . . . .	6
3.1 Processo do mapeamento sistemático . . . . .	14
3.2 Processo de triagem de estudos primários . . . . .	19
4.1 Número de estudos primários mapeados e ano de publicação . . . . .	28
4.2 Resultado da questão de pesquisa 2 - Quantidade de ferramentas e fases suportadas . . . . .	31
4.3 Resultado da questão de pesquisa 2.1 - Quantidade de ferramentas por nível de abstração da linguagem de especificação para cada fase suportada . . . . .	32
4.4 Resultado da questão de pesquisa 2.2 - Quantidade de ferramentas por nível de automação para cada fase suportada . . . . .	34
4.5 Resultado da questão de pesquisa 3 - Quantidade de ferramentas por tipo de experimento suportado . . . . .	37
4.6 Fases suportadas, nível de automação e nível de abstração de linguagem para ferramentas de apoio a experimentos Não Baseados em Tecnologia . . . . .	38
4.7 Fases suportadas, nível de automação e nível de abstração de linguagem para ferramentas de apoio a experimentos Baseados em Tecnologia . . . . .	40
4.8 Fluxograma do guia preliminar para seleção de ferramenta de apoio a um experimento . . . . .	47

# Lista de Tabelas

3.1	Questões de pesquisa . . . . .	15
3.2	Veículos de publicação da busca manual . . . . .	17
3.3	CrITÉrios de incluso e de excluso de estudos primrios . . . . .	18
4.1	Veículos de publicação, perÍodo pesquisado e quantidade de estudos . . . . .	23
4.2	Estudos primrios e ferramentas . . . . .	24
4.3	Veículos de publicação e quantidade de estudos primrios selecionados . . . . .	26
4.4	Ferramentas, fases de projeto de experimento suportadas, nÍvel de automa- ço e nÍvel de abstraço de linguagem . . . . .	27
4.5	Resultado da questo de pesquisa 1 - Lista de ferramentas . . . . .	30
4.6	Questionrio para identificaço das necessidades de ferramenta de apoio a experimentos . . . . .	43
4.7	Apoio oferecido pelas ferramentas no baseadas em tecnologia . . . . .	49
4.8	Apoio oferecido pelas ferramentas baseadas em tecnologia . . . . .	50
5.1	Questes de pesquisa X trabalhos relacionados . . . . .	54

# Capítulo 1

## Introdução

Realizar experimentos para obter resultados em estudos científicos, de forma controlada, já é utilizado em várias áreas há muitos anos, mas na Engenharia de Software é aplicada há pouco tempo, se comparada a outras. A execução de experimentos em pesquisas demonstra rigor em seu conteúdo, pois analisa empiricamente cada resultado obtido pela alteração de determinadas variáveis, e assim oferece maior confiabilidade em sua proposta comparado a estudos que não comprovam a eficácia ou eficiência de seus métodos.

Para uma melhor organização e desenvolvimento de experimentos em Engenharia de Software, Wohlin et al. (2012) e Juristo e Moreno (2013) definiram estratégias com passos bem definidos para sua execução e coleta de dados. Executar um projeto de experimento pode exigir muito esforço dos pesquisadores em Engenharia de Software, e ferramentas foram desenvolvidas para apoiar nesse processo.

Para identificar e atualizar as opções de ferramentas de apoio a experimentos existentes, a metodologia para realizar um mapeamento sistemático, conforme definido por Kitchenham e Charters (2007), pode ser executada. Dessa forma é possível obter, de forma estruturada, informações que possam auxiliar a pesquisadores em Engenharia de Software na escolha de uma ferramenta, quando necessitarem elaborar um experimento controlado. Uma revisão sistemática realizada por Freire et al. (2013) já identificou algumas dessas ferramentas, e levantou propriedades sobre como e quais fases de um projeto de experimentação são suportadas.

O objetivo desse trabalho é fazer um mapeamento sistemático mais atual e abrangente sobre ferramentas de apoio a experimentos em Engenharia de Software, e que poderá ser útil a pesquisadores que buscam uma ferramenta para realizar experimentos, pois são apresentados pontos pelos quais facilitará escolher uma dentre as opções encontradas, seja de acordo com a cobertura de apoio às fases de um projeto de experimentos, ou por outro item mapeado por nossas questões de pesquisa.

## 1.1 Definição do Problema

Projetos propostos pela academia e pela indústria necessitam de uma validação por técnicas e métodos de estudos empíricos para que tenham mais aceitação pela comunidade. Dentre estes métodos está a execução de experimentos controlados, seguindo o guia proposto por Wohlin et al. (2012), com diversas instruções das etapas que devem ser cumpridas para que um projeto realize um experimento de forma apropriada.

Quando um pesquisador em Engenharia de Software necessita realizar um experimento sobre seu projeto, pode optar por construir manualmente todos os recursos exigidos para a execução do experimento, que, dentre eles, estão as tarefas de planejar, executar e coletar dados para utilizá-los em uma análise estatística. Outra opção, com intuito de guiar a operacionalização de todas essas tarefas para evitar erros, é de utilizar uma ferramenta de apoio para conduzir experimentos de uma forma mais correta.

O problema tratado neste trabalho é determinar como os pesquisadores em Engenharia de Software podem obter uma lista com as ferramentas de apoio a experimentos existentes, além de identificar parâmetros e funcionalidades que os auxiliem na escolha de qual opção melhor se encaixa no experimento que desejam produzir.

A relevância desse problema é que, sem um mapeamento do estado da arte, a busca por ferramentas existentes pode não retornar os resultados esperados, exigir muito tempo para encontrá-las, ou até ocasionar retrabalho, quando pesquisadores em Engenharia de Software desenvolvem um sistema próprio para controlar seus projetos de experimentos.

Uma lista de ferramentas que oferecem suporte a experimentos em engenharia de software foi publicada pela revisão sistemática de Freire et al. (2013), que apresentou resultados obtidos até dezembro de 2011. No entanto, desde então, não foi encontrada uma nova pesquisa com a mesma abrangência que a revisão de Freire et al. (2013).

Além disso, outras informações podem ser extraídas quanto às fases de um projeto de experimentação, como identificar na ferramenta o grau de abstração da linguagem utilizada na especificação de experimentos, o nível de automação suportado, a sua confiabilidade e se suporta experimentos baseados em tecnologia. Essas informações adicionais auxiliariam o pesquisador em Engenharia de Software na escolha de uma determinada ferramenta, pois são indicativos de qualidade e produtividade.

## 1.2 Solução Proposta

Como objetivo deste trabalho realizamos um mapeamento sistemático, com intenção de sintetizar a evidência sobre ferramentas de apoio a experimentação em Engenharia de Software existentes. Além de realizar esse levantamento, também elaboramos uma lista

de atributos para segmentar as ferramentas em tipos de utilização, que se tornaram as questões de pesquisa desta dissertação.

O produto esperado deste trabalho é identificar o estado da arte sobre ferramentas de apoio a experimentos, e que resolva o problema citado anteriormente, apoiando os pesquisadores em Engenharia de Software quando necessitarem escolher uma ferramenta de suporte a experimentos.

Para obter esse resultado, seguimos o guia de Kitchenham e Charters (2007) para elaboração do protocolo de mapeamento sistemático. Com isso determinamos os parâmetros necessários para realizar a pesquisa em estudos já publicados, e selecionar dentre eles os que atendessem aos nossos requisitos.

No final deste mapeamento, obtivemos a visão geral sobre as técnicas utilizadas pelas ferramentas, assim como quais etapas de experimentação elas compreendem, se algum processo de automação é oferecido, que tipo de linguagem é utilizada para especificação de um experimento, qual a confiabilidade da ferramenta é obtida pela garantia proposta pelos autores e se suporta experimentos baseados em tecnologia. Também identificamos neste trabalho os problemas mais comuns encontrados nos estudos primários, ou pontos que simplesmente não foram tratados em suas propostas.

### **1.3 Resumo das Contribuições**

Neste trabalho, apresentamos as seguintes contribuições:

- Identificação de sete ferramentas de apoio a experimentos, a definição das fases de um projeto de experimentos suportadas, o nível de automação e grau de abstração da linguagem de especificação utilizada;
- Análise sobre as ferramentas de apoio a experimentos;
- Levantamento dos problemas mais comuns entre as ferramentas disponíveis, com recomendações para melhorias e sugestões para novas ferramentas;
- Elaboração de um guia preliminar para auxílio dos pesquisadores em Engenharia de Software na escolha de uma ferramenta de apoio a experimentos.

### **1.4 Organização do Trabalho**

Este trabalho está organizado da seguinte forma:

O Capítulo 2 apresenta o referencial teórico sobre o assunto que dissertamos. Introduzimos a metodologia para realizar experimentos em engenharia de software de Wohlin et al.

(2012), uma breve descrição dos trabalhos que utilizamos como referência e o processo seguido para elaborar o protocolo de pesquisa.

O Capítulo 3 detalha a metodologia de pesquisa para realizar o mapeamento sistemático, seguindo o guia definido por Kitchenham e Charters (2007). No protocolo está definida a estratégia de busca de estudos primários, o critério de seleção, o processo de triagem realizado, a forma de extração dos dados e a como foi realizada a análise dos dados. Esse capítulo também apresenta as questões de pesquisa que utilizamos neste trabalho.

No Capítulo 4 são apresentados e analisados os resultados obtidos pelo mapeamento sistemático de forma geral, mas também uma análise individual respondendo a cada uma das questões de pesquisa, além de identificar pontos de falha que não são cobertos pelas ferramentas. Também é apresentado neste capítulo um guia preliminar para auxílio na escolha da ferramenta mais adequada, para um pesquisador em Engenharia de Software que deseje conduzir um experimento com apoio de ferramentas.

O Capítulo 5 apresentamos a conclusão do nosso estudo, discutindo trabalhos futuros que podem ser realizados baseados nos resultados do nosso mapeamento, assim como nossas conclusões e ameaças à validade do trabalho realizado. Este capítulo também descreve em detalhes os trabalhos relacionados e já publicados com o mesmo tema que nossa pesquisa. Apresentamos as diferenças e o valor agregado esperado com nossa proposta, comparando tanto as questões de pesquisa, os veículos de publicação usados na busca, assim como o período compreendido entre os trabalhos.

O Anexo I contém a versão completa dos formulários de extração de dados, utilizados para obter as informações sobre as ferramentas de apoio a experimentos encontradas nos estudos primários. Por fim, o Anexo II apresenta os formulários de extração preenchidos para cada ferramenta identificada.

# Capítulo 2

## Referencial Teórico

Neste capítulo apresentamos os conceitos que foram utilizados para elaboração desse trabalho. A introdução ao conceito de experimentos em engenharia de software, na Seção 2.1, e a metodologia para realizar o mapeamento sistemático, na Seção 2.2.

### 2.1 Experimentos em Engenharia de Software

Existem várias formas de realizar uma pesquisa científica, porém nenhuma delas pode ser baseada apenas em opiniões ou interesses comerciais, conforme definido por Juristo e Moreno (2001). Dentre os formatos de estudos empíricos podemos citar surveys, revisões sistemáticas, estudos de caso, e experimentos controlados.

Experimentos são os estudos realizados com maior controle, onde diferentes tratamentos são aplicados a diferentes sujeitos, mantendo algumas variáveis constantes, e medindo os resultados de outras. Suas principais vantagens são o controle de indivíduos, objetos e instrumentação. Outra vantagem é a habilidade para realizar análises estatísticas, além de oferecer oportunidades para replicação dos experimentos.

Para avaliar nossas convicções, podemos usar um experimento, que é criado, por exemplo, para testar uma teoria ou hipótese. O ponto inicial é a idéia de um relacionamento causa e efeito, que podemos formalizar em uma hipótese.

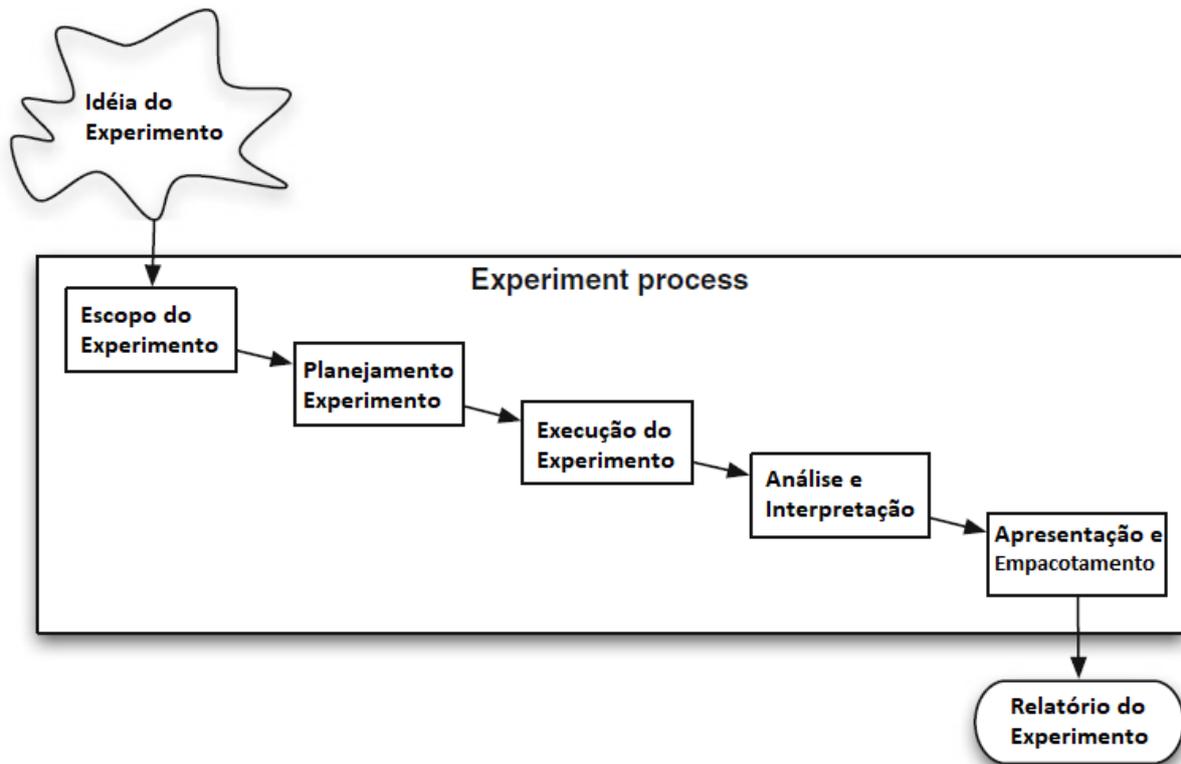
Existem experimentos baseados em tecnologia, onde diferentes tratamentos técnicos são aplicados a diferentes objetos, como, por exemplo, duas ferramentas de teste são aplicadas ao mesmo código-fonte. Já os experimentos que não são baseados em tecnologia, humanos aplicam esses tratamentos, como, por exemplo, dois programadores codificam sistemas com as mesmas funcionalidade, e posteriormente comparam a qualidade dos códigos desenvolvidos por ambos. Experimentos baseados em tecnologia tem um controle maior, pois tendem a ser determinísticos, diminuindo a chance de um erro humano.

O assunto principal deste trabalho trata sobre ferramentas que apóiam a realização de experimentos na forma de estudos empíricos. Existem estudos que abordam a metodologia de produzir e conduzir experimentos em Engenharia de Software, como o de Wohlin et al. (2012), que utilizaremos como referência, e o de Juristo e Moreno (2013). Optamos neste trabalho seguir a proposta de Wohlin et al. (2012) porque os pesquisadores já tinham experiência nesta metodologia.

Para definir um processo de experimentação, Wohlin et al. (2012) aborda as fases de escopo, planejamento, execução, análise e interpretação, apresentação e empacotamento. Com base nessas fases que dividimos a cobertura de apoio oferecido pelas ferramentas pesquisadas em nosso mapeamento.

A Figura 2.1 apresenta o fluxo entre as fases de um projeto de experimentos.

Figura 2.1: Fases de um projeto de experimentos



O detalhamento das cinco fases do processo de experimento são detalhadas nas Seções 2.1.1 sobre Escopo, 2.1.2 sobre Planejamento, 2.1.3 sobre Execução, 2.1.4 sobre Análise e Interpretação, e, por último, 2.1.5 sobre Apresentação e Empacotamento.

### 2.1.1 Fase de Escopo

Na fase de escopo é determinada a fundamentação do experimento, o motivo a ser conduzido. Se tiver corretamente elaborado, o experimento pode ser usado para estudar o que foi pretendido.

O escopo do experimento é realizado definindo seus objetivos. O modelo definido por Wohlin et al. (2012) para elaboração de um objetivo é: Analisar <Objeto(s) de estudo> para o propósito de <Propósito> com respeito a seus <Foco da qualidade> do ponto de vista da <Perspectiva> no contexto de <Contexto>.

Entende-se com Objeto a entidade que é estudada no experimento, que podem ser produtos, processos, recursos, modelos, métricas ou teorias. O Propósito é a definição da intenção do experimento, como avaliar o impacto de duas técnicas de programação diferentes, por exemplo.

O Foco da Qualidade descreve o efeito que está tentando buscar com o experimento, como eficácia, custo, confiabilidade, etc. A Perspectiva indica o ponto de vista do qual os resultados do experimento devem ser interpretados. O Contexto define quais indivíduos (pessoas ou tecnologias) estão envolvidas, e quais objetos (artefatos de software) serão utilizados no experimento. Os indivíduos definidos pelo Contexto é que aplicarão os tratamentos nos objetos descritos, durante a execução do experimento.

Esse modelo, quando preenchido, forma a definição do objetivo do experimento, que é a entrada para o passo de planejamento no processo de experimento.

### 2.1.2 Fase de Planejamento

Conforme definição de Wohlin et al. (2012), a fase de Escopo determina a fundação para o experimento, o motivo pelo qual o experimento é conduzido, enquanto a fase de Planejamento prepara a forma como o experimento é conduzido.

A fase de Planejamento é dividida nos seguintes passos: seleção do contexto, formulação da hipótese, seleção de variáveis, seleção de indivíduos, projeto do experimento, instrumentação e avaliação da validade.

Durante o passo de Seleção de Contexto, é definido o ambiente em qual o experimento será executado. Projetos de software reais de uma empresa, ou testes simulados por estudantes, são formas de contexto. cada uma tem suas vantagens e desvantagens, e fica a cargo do experimentador definir qual é mais interessante para seu projeto.

A Formulação da Hipótese gera a base para realizar a análise estatística do experimento, pois realizará testes nas hipóteses definidas neste passo, utilizando os dados coletados após sua execução. A hipótese nula é a que o experimentador deseja que seja rejeitada, e uma hipótese alternativa é a que deve ser contrária ao da hipótese nula. Testes estatísticos

devem ser escolhidos de acordo com as hipóteses, e serão utilizados para avaliar o resultado de um experimento.

O passo de Seleção de Variáveis é onde serão definidas as variáveis dependentes e as independentes. As independentes são aquelas variáveis que podem ser controladas pelo experimentador e alteradas durante a execução dos experimentos. As independentes devem ter escalas de medidas definidas, assim como os limites de abrangência e os níveis específicos que os testes serão realizados. As variáveis dependentes são aquelas que serão afetadas pelas independentes, e que os efeitos dos tratamentos aplicados serão medidos. As dependentes são derivadas diretamente da hipótese do experimento.

Na Seleção de Indivíduos é definida como uma amostra da população alvo do experimento, podendo ser probabilística, quando são escolhidos os indivíduos, ou não-probabilística, quando a escolha de indivíduos é aleatória. O tamanho da amostra também define a precisão dos resultados de um experimento. Quanto maior a variabilidade da população, uma amostra maior pode ser necessária.

No passo de Projeto do Experimento, para obter conclusões significativas de um experimento, aplicamos métodos de análise estatística nos dados coletados para interpretar os resultados. Para o projeto, devemos saber qual análise estatística executar para rejeitar a hipótese nula. Logo, o projeto e a interpretação dos resultados estão muito relacionados. Também devem ser determinados quantos testes o experimento deve ter, para certificar que o efeito do tratamento é visível.

Ainda na elaboração do Projeto, devem ser definidos como serão alocados os objetos, os indivíduos e em qual ordem os testes serão executados. Para isso podem ser aplicados princípios gerais de projetos, como aleatorização, bloqueamento e balanceamento. Wohlin et al. (2012) cita padrões de projeto para experimentos, para serem escolhidos de acordo com o número de fatores e de tratamentos a que os experimentos serão submetidos.

O passo de Instrumentação tem como objetivo definir quais instrumentos serão escolhidos ou desenvolvidos para um experimento ser executado. Nesse passo os objetos devem ser selecionados, como, por exemplo, documentos de especificação ou de códigos. Também podem ser definidos guias para auxiliar os participantes, como listas de verificação e outros materiais de apoio. Instrumentos para realizar a coleta de dados durante a execução do experimento devem ser definidos, como formulários para preenchimento ou ferramentas para captura das informações. Esses instrumentos devem prover meios para executar o experimento e monitorá-lo, sem afetar seu controle e seu resultado.

A Avaliação da Validade do experimento deve constar no projeto, determinando a validade dos resultados obtidos. Primeiro se observa a validade para a população de qual a amostra foi selecionada para o experimento, e posteriormente se também pode ser aplicada a uma população mais ampla, caso essa generalização for de interesse do

experimentador. Ameaças à validade de um experimento são definidas em quatro tipos, por Cook e Campbell (1979), sendo elas conclusão, interna, de construção ou externa.

### **2.1.3 Fase de Execução**

Após o experimento ser definido e planejado, ele deve ser executado para que dados sejam gerados para posterior análise. É nesta fase operacional que os tratamentos são executados nos objetos pelos indivíduos, reiterando que indivíduos podem ser pessoas ou baseados em tecnologia, como máquinas e sistemas.

Os principais passos para esta fase operacional do experimento são a preparação, a execução e a validação de dados. Na preparação os indivíduos são escolhidos, assim como outros fatores necessários para a realização do experimento. A execução são as tarefas realizadas pelos indivíduos a diferentes tratamentos, gerando dados sobre os experimentos, que em seguida são submetidos para o passo de validação de dados.

O passo de Preparação organiza o experimento para ser executado. Quanto mais organizado, mais fácil será em executá-lo. Este passo consiste basicamente em selecionar os indivíduos que participarão do experimento, assim como separar os materiais necessários para a execução, como formulários e ferramentas.

Para selecionar os indivíduos participantes de um experimento que não é baseado em tecnologia, assim executado por pessoas, deve-se levar em consideração o conhecimento daquele indivíduo quanto ao experimento que ele fará parte. Caso o experimento esteja fora do contexto daquele indivíduo, pode ser uma ameaça a validade do mesmo, quando este tiver que executar uma atividade estranha a ele.

A separação dos instrumentos necessários para a execução do experimento também faz parte da Preparação. Neste passo os objetos que serão submetidos ao experimento, os guias, formulários e ferramentas, devem estar prontos para uso, assim como determinados na fase de Planejamento. Pode ocorrer que, para cada tratamento a ser executado pelo experimento, tenham instrumentos diferentes a serem utilizados.

O passo de Execução, seguindo o que foi definido na fase de Planejamento, pode ser realizado de várias formas, como, por exemplo, em um processo único em que todos os indivíduos estão reunidos e executam os tratamentos do experimento ao mesmo tempo, como também um experimento que necessita de um grande espaço de tempo para ser concluído.

A coleta dos dados executados no experimento também é realizada neste passo, podendo ser realizada pelos próprios indivíduos, em formulários, ou suportados por ferramentas. No caso de ferramentas apoiarem este passo, é possível que sejam apenas um instrumento para terem os dados preenchidos manualmente pelos indivíduos, mas também podem realizar a coleta de dados de forma totalmente automática.

O passo final desta fase é a Validação de Dados, realizada pelo experimentador para verificar se o experimento foi realizado corretamente pelos indivíduos, que os tratamentos foram aplicados na forma e ordem planejada, e os dados foram coletados de corretamente.

#### **2.1.4 Fase de Análise e Interpretação**

Após realizada a fase de Execução do experimento, dados foram coletados e é o momento para analisá-los e interpretá-los para gerar informações de interesse dos experimentadores. Essa análise quantitativa é realizada em três passos, sendo eles: efetuar uma estatística descritiva, reduzir o conjunto de dados e testar a hipótese definida para o experimento.

O passo de Estatística Descritiva realiza o processamento numérico, utilizando métricas bem definidas, em um conjunto de dados para obter descrições e gráficos interessantes sobre aquele aspecto. Com isso é possível analisar como os dados estão distribuídos, e também identificar dados anormais e pontos fora da curva.

A seguir, o passo de Redução do Conjunto de Dados é aplicado, caso os dados trabalhados pelos métodos estatísticos não representarem o que deveriam, e assim determinando que as conclusões obtidas foram levadas ao erro. Esse problema pode ocorrer por erros sistemáticos nos conjuntos de dados, ou por pontos fora da curva, que podem ser detectados por métodos estatísticos.

Quando detectados problemas dessa forma no conjunto de dados, é necessário decidir como proceder, se aplicar uma redução na massa de dados para excluir um ponto fora da curva que não é interessante ao estudo, ou se o deixa a vista para futura análise pelo experimentador.

Por fim, o passo de Teste de Hipótese é o último a ser executado, para validar se é possível rejeitar uma hipótese nula, baseado em alguns dados estatísticos obtidos nos passos anteriores. O experimentador seleciona uma amostra dos dados e tenta rejeitar que aquelas propriedades sejam verdadeiras. Quando encerradas essas validações, o experimentador descreve as conclusões sobre a influência, ou não, das variáveis independentes sobre as dependentes. Com isso é possível provar se há uma diferença estatística significativa entre os tratamentos aplicados no experimento.

#### **2.1.5 Fase de Apresentação e Empacotamento**

Assim que o experimento é concluído, há necessidade de apresentar os resultados e, se possível, empacotar todo o material produzido para futura replicação por outras pessoas. Artigos ou relatórios técnicos podem ser produzidos para a comunidade, ou os dados dos experimentos podem ser armazenados em uma base de dados corporativa, dependendo da finalidade a que foi submetido.

Empacotar experimentos com finalidade de replicação permite que outros experimentadores possam verificar as propriedades como o experimento foi realizado, e a forma como os dados foram obtidos por sua execução. Ambientes computacionais inteiros podem ser adicionados nesse pacote, quando o tipo de experimento permitir, como quando utilizado recursos como máquinas virtuais e similares.

Há possibilidade de experimentos empacotados sejam customizados, com outras variáveis, ou pela execução por outros indivíduos, ou por outras formas de tratamento, para realizarem novos experimentos, mas servindo como uma base inicial de projeto. Isso é mais provável de ocorrer em ferramentas que apóiem experimentos baseados em tecnologia, que tem pouca interferência humana.

Para a apresentação do relatório do experimento em um formato acadêmico, Jedlitschka e Pfahl (2005) propõe um modelo dividido de forma estruturada em doze itens, para produzir um material com maior número de informações sobre o experimento realizado.

## **2.2 Processo de Revisão da Literatura**

Pela definição de Kitchenham e Charters (2007), uma revisão sistemática da literatura é uma forma de identificar, avaliar e interpretar todas as pesquisas relevantes de uma questão de pesquisa em particular, ou área de um tópico, ou um fenômeno de interesse. Uma revisão sistemática é uma forma de estudo secundário, onde os estudos que forem selecionados através dessa pesquisa são chamados de estudos primários.

Alguma forma de revisão de literatura é utilizada na maioria das pesquisas, com objetivo de identificar lacunas em estudos existentes, resumir as evidências, benefícios e limitações sobre alguma tecnologia ou metodologia, identificar novas áreas para investigação, dentre outros motivos. Mas para que esse processo tenha um valor científico maior, é sugerido que seja realizada uma revisão de forma sistemática, que esteja de acordo com alguma estratégia de pesquisa predefinida.

Dentre as vantagens de uma revisão sistemática da literatura podemos citar que é uma metodologia bem definida que demonstra um viés menor nos resultados obtidos, também pode prover informação sobre os efeitos de algum fenômeno específico sobre uma grande abrangência de métodos empíricos, assim como em estudos quantitativos é possível combinar dados utilizando técnicas analíticas, que aumentam a probabilidade de detectar efeitos reais, que em estudos menores não são possíveis de serem identificados.

### **2.2.1 Mapeamento Sistemático**

Mapeamentos sistemáticos são revisões da literatura similares às revisões sistemáticas, e tanto Kitchenham e Charters (2007) como Petersen et al. (2008) descrevem suas diferenças.

Dentre elas, está a definição das questões de pesquisa, que em um mapeamento são mais abrangentes, para obter um número maior de estudos primários, mas que se aprofundam na classificação e categorização dos resultados. Em uma revisão sistemática, o foco é maior nos detalhes de cada estudo primário. A fase de análise de um mapeamento também é mais ampla do que em uma revisão, em geral sumarizando e totalizando os resultados ao invés de aplicar técnicas de análise muito profundas.

Como referência para elaboração do protocolo de pesquisa deste mapeamento sistemático utilizamos o guia de Kitchenham e Charters (2007). Nele é especificado todas as etapas para a realização de revisões e mapeamentos sistemáticos. Também utilizamos o estudo de Vale et al. (2017), que apresentaram uma aplicação dessa metodologia para execução de um mapeamento sistemático, e um modelo para o protocolo de pesquisa. Estudamos também a proposta de Petersen et al. (2008) para realização de mapeamentos sistemáticos. As duas propostas se correspondem, com o trabalho de Petersen et al. (2008) apresentando mais detalhes na condução de mapeamentos sistemáticos, mas optamos pelo guia de Kitchenham e Charters (2007) por ser mais consolidado, atendendo as especificações necessárias para nosso estudo, além de que já temos experiência por ter aplicado esse guia no estudo de Vale et al. (2017).

No Capítulo 3 apresentamos a descrição das etapas da elaboração do mapeamento sistemático, com detalhamento das suas atividades e a aplicação direcionada às nossas questões de pesquisa.

# Capítulo 3

## Método de Pesquisa

Para identificar as ferramentas de apoio à experimentos existentes e todas as propriedades definidas neste estudo, a metodologia que utilizamos foi a realização de um mapeamento sistemático seguindo o guia de Kitchenham e Charters (2007), assim como a adaptação produzida por Vale et al. (2017). A escolha do mapeamento ao invés de uma revisão sistemática se dá pelo fato que esta última tem um escopo mais reduzido, geralmente direcionado a uma determinada área específica e com maior detalhamento, além disso, trabalhos preliminares indicam uma baixa qualidade dos estudos primários, e um mapeamento usualmente não avalia a qualidade dos estudo.

Através de um mapeamento sistemático, é possível representar o estado da arte e classificar os resultados de forma mais abrangente, identificando o grau de aderência das ferramentas a projetos de experimentação. Também pretendemos com esse estudo identificar os problemas mais comuns entre as ferramentas e as suas principais funcionalidades.

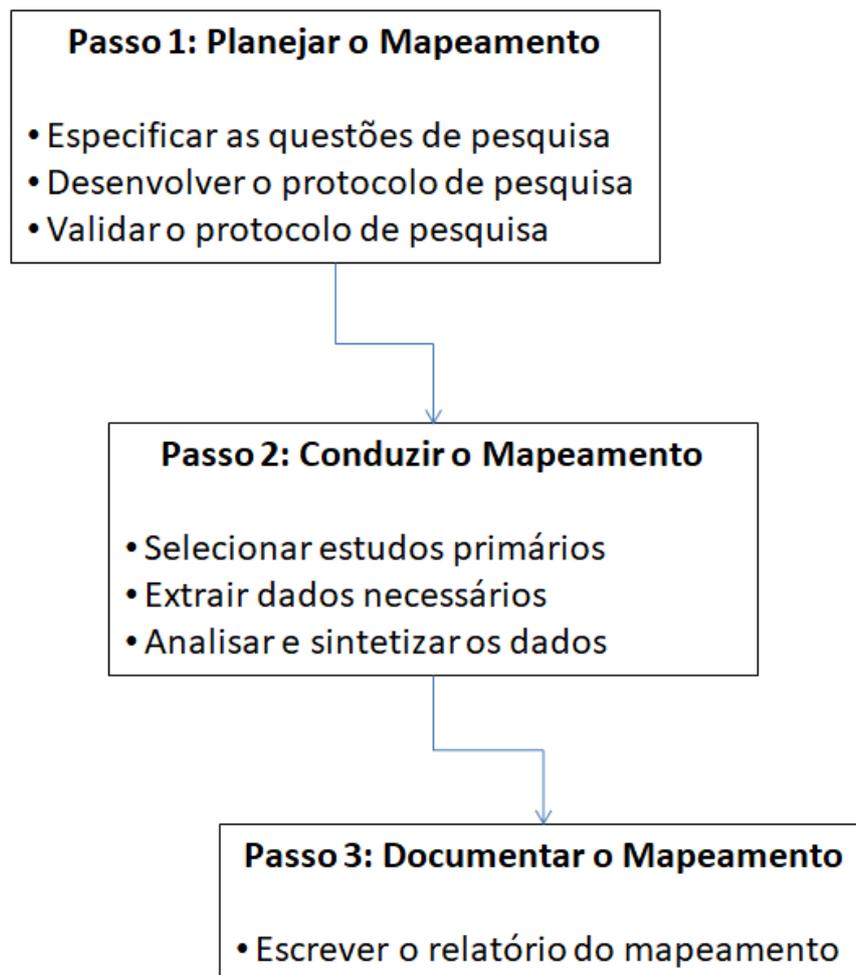
Na Figura 3.1 estão identificados os principais passos e o fluxo do processo de mapeamento sistemático que utilizamos neste trabalho. O primeiro passo é o planejamento do mapeamento sistemático, que trata as questões essenciais do trabalho, como mapear as necessidades de informação em questões de pesquisa, elaborar o protocolo que deverá ser seguido para sua execução e validá-lo realizando uma pesquisa inicial, para verificar se os resultados obtidos atendem às questões de pesquisa definidas. O segundo passo trata da execução do mapeamento conforme o protocolo de pesquisa especificado, com objetivo de extrair e analisar os dados obtidos. O último passo é a documentação o resultado do mapeamento.

Para planejar este mapeamento sistemático, elaboramos três questões de pesquisa, sendo que a segunda tem mais três subquestões. Desenvolvemos o protocolo de pesquisa, apresentado neste capítulo, que foi validado parcialmente aplicando os critérios de seleção e realizando a extração dos dados em dois estudos primários, já identificados na revisão sistemática de Freire et al. (2013), que apresentam ferramentas de apoio a experimentos.

Nessa validação, verificamos que os critérios utilizados para extração dos dados conseguiram atender as nossas questões de pesquisa, e assim estudos que proponham ferramentas de apoio a experimentos em engenharia de software serão identificados pela nossa busca, e terão suas informações extraídas corretamente.

No passo de condução do experimento, realizamos a seleção de estudos primários de acordo com os critérios definidos na Seção 3.3, extraímos os dados desses estudos e efetuamos a síntese sobre as informações levantadas no Capítulo 4, de Resultados e Análise.

Figura 3.1: Processo do mapeamento sistemático



Neste capítulo, apresentamos a Seção 3.1, que descreve as nossas questões de pesquisa, a Seção 3.2 com a estratégia de busca, a Seção 3.3 com os critérios de seleção, a Seção 3.4 com o processo de triagem dos estudos, a Seção 3.5 com o procedimento para a extração dos dados, e a Seção 3.6, que descreve como será realizada a análise e síntese dos dados extraídos.

## 3.1 Questões de Pesquisa

As questões de pesquisa estão detalhadas na Tabela 3.1, juntamente com a motivação que nos levou a cada uma delas.

Tabela 3.1: Questões de pesquisa

#	Questão de Pesquisa	Motivação
QP1	Quais são as ferramentas de apoio à experimentação em Engenharia de Software e as funcionalidades de experimentação suportadas pelas ferramentas?	O objetivo desta questão é identificar quais são as ferramentas existentes, levantando os dados básicos como nome, sua forma de execução (funcionamento) e quais as características principais para apoio no desenvolvimento de um projeto de experimentação. Fornece-nos as informações básicas para desenvolver este mapeamento sistemático.
QP2	Quais são as fases de um projeto de experimentação cobertas pela ferramenta?	As fases de um projeto com experimentos são divididas entre: escopo, planejamento, execução, análise e interpretação, apresentação e empacotamento. Com essa informação é possível definir se as ferramentas oferecem suporte a uma ou mais fases, e assim o grau de aderência a um projeto de experimentação.
QP2.1	Qual o nível de abstração da linguagem de especificação da ferramenta?	Na modelagem de um experimento, a ferramenta pode apoiar cada fase em um nível de abstração diferente, com maior ou menor detalhamento nas suas definições sintáticas e semânticas. Um nível alto de abstração seria utilizar linguagem natural na especificação de um experimento, usando diretamente conceitos de hipóteses, tratamentos, objetos, etc. Já um nível de abstração mais baixo seria a utilização de uma linguagem de scripts, como as utilizadas em linguagem de programação. Identificar o nível de abstração da linguagem, para cada fase do projeto de experimentos, permite identificar sua aderência ao domínio de experimentação, e a eventual necessidade adicional de conhecimento de programação..

QP2.2	Qual é o nível de automação suportado pela ferramenta?	Durante a fases de um experimento, é possível que a ferramenta forneça suporte em três níveis: automatizado, executando o experimento sem necessidade de intervenção humana; semiautomatizado, quando há uma intervenção mínima; e manual. Com essa categorização, pretendemos identificar o nível de automação que a maioria das ferramentas de apoio a experimentos oferecem para cada fase, e assim se um trabalho futuro pode ser realizado para facilitar a implementação de automação nos experimentos.
QP2.3	Qual é a confiabilidade fornecida pela ferramenta?	Em cada fase coberta pela ferramenta, é esperado que seja definida uma garantia aos seus usuários de que o suporte oferecido por ela não produzirá erros no processo de experimentação. Pretendemos identificar o nível de confiabilidade proposto para cada ferramenta, de acordo com as garantias apresentadas nos estudos.
QP3	Que tipo de experimento é apoiado pela ferramenta?	Consideramos neste estudo dois tipos de experimentos: aqueles que utilizam tecnologia para sua execução, onde os tratamentos aplicados aos objetos do experimento são executados por programas, e os que são realizados diretamente por pessoas. Nosso objetivo com essa questão é classificar os estudos primários e correlacionar os resultados com as outras questões de pesquisa, a fim de identificar oportunidades de aprimoramento das ferramentas.

## 3.2 Estratégia de busca

A estratégia de busca que utilizamos para selecionar os estudos primários foi a pesquisa manual, aplicando nos principais veículos de publicações com temas relacionados ao assunto do nosso trabalho.

A pesquisa automatizada foi descartada, pois em uma verificação inicial, utilizando palavras chave como "experiment", "project", "tool", "software", dentre outras, foram retornados milhares de resultados, pois várias publicações tem conteúdos com essas palavras, mas não especificamente com ferramentas de apoio a experimentação. Dessa forma, seria necessário realizar a pesquisa manual nos resultados dessa pesquisa automatizada, para aplicar os critérios de seleção estudo por estudo, e não incluir um item fora do escopo em

nosso mapeamento. Então optamos por realizar apenas a busca manual, por ter um grau de precisão mais elevado, conforme previamente observado por Kitchenham et al. (2010).

Os veículos de publicação selecionados para a pesquisa foram periódicos, conferências, e workshops relevantes ao objetivo do nosso mapeamento, e estão descritos na Tabela 3.2.

Tabela 3.2: Veículos de publicação da busca manual

<b>Veículo de Publicação</b>
IEEE Transactions on Software Engineering (TSE)
ACM Transactions on Software Engineering and Methodology (TOSEM)
Journal on Software and Systems Modeling (SoSyM)
International Conference on Automated Software Engineering (ASE)
European Software Engineering Conference and the Symposium on the Foundations of Software Engineering (ESEC/FSE)
Journal of Empirical Software Engineering (ESEJ)
Journal of Systems and Software (JSS)
ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)
Experimental Software Engineering Latin American Workshop (ESELAW)
International Conference on Software Engineering (ICSE)

A pesquisa manual também foi realizada em estudos similares ao nosso, aplicando nossas questões de pesquisa aos estudos primários já identificados, conforme descrito no Capítulo 5. Na Tabela 4.1, estão apresentados os períodos compreendidos para cada veículo de publicação utilizado na busca e a quantidade de estudos pesquisados. Cinco veículos de publicação foram utilizados no estudo de Freire et al. (2013), compreendendo o período de janeiro de 2002 até dezembro de 2011, e nesses veículos efetuamos a busca atualizando até a janeiro de 2018. Também incluímos outros cinco veículos compreendendo o período desde janeiro de 2002 até janeiro de 2018.

### 3.3 Critério de Seleção

Para identificar estudos primários relacionados às nossas questões de pesquisa, definimos critérios de inclusão e de exclusão dos estudos primários. O interesse do nosso trabalho é encontrar estudos que apresentem ferramentas de apoio a experimentos em engenharia de software, logo este é o primeiro critério de inclusão.

O segundo critério de inclusão utilizado é que os estudos primários estivessem dentre os veículos de publicação descritos na Tabela 3.2, ou na revisão sistemática realizada por Freire et al. (2013). Adicionalmente, selecionamos apenas artigos escritos em língua

inglesa, com intuito em abranger um número maior de experimentadores que desejem ler os estudos primários que apresentam as ferramentas mapeadas.

Um critério de exclusão foi remover estudos duplicados, que podem ter sido publicados em diferentes veículos mas que detenham o mesmo conteúdo. Aquela publicação com mais detalhes que será levada em consideração. O segundo critério de exclusão é excluir os estudos primários que apresentem apenas frameworks de experimentação, e não propriamente ferramentas que possam ser utilizadas pelos pesquisadores em Engenharia de Software.

Alguns estudos primários apenas reportavam frameworks para experimentação, porém não apresentavam um software como ferramenta para utilização desses frameworks, logo estes estudos não se enquadraram em nosso mapeamento sistemático.

Estudos primários com informações insuficientes sobre a ferramenta de experimentação, como os que apresentassem apenas o nome do produto, mas sem citar nenhuma de suas funcionalidades, também foram excluídos. A Tabela 3.3 apresenta os critérios de inclusão e de exclusão.

Tabela 3.3: Critérios de inclusão e de exclusão de estudos primários

<b>Critérios de Inclusão</b>	
01	Estudos que apresentem ferramentas de apoio a experimentos em engenharia de software
02	Estudos pertencentes aos veículos de publicação selecionados
03	Estudos escritos na língua inglesa
<b>Critérios de Exclusão</b>	
01	Estudos duplicados
02	Estudos que apresentem apenas frameworks de experimentação
03	Estudos com informações insuficientes sobre a ferramenta
04	Estudos fora do escopo deste mapeamento sistemático

### 3.4 Processo de triagem

O processo de triagem é a forma como os estudos são filtrados pelos pesquisadores. Os estudos primários identificados pelo mapeamento são obtidos quando aplicados os critérios de inclusão, para criar a lista inicial de estudos selecionados, e logo após remover itens dessa lista aplicando os critérios de exclusão.

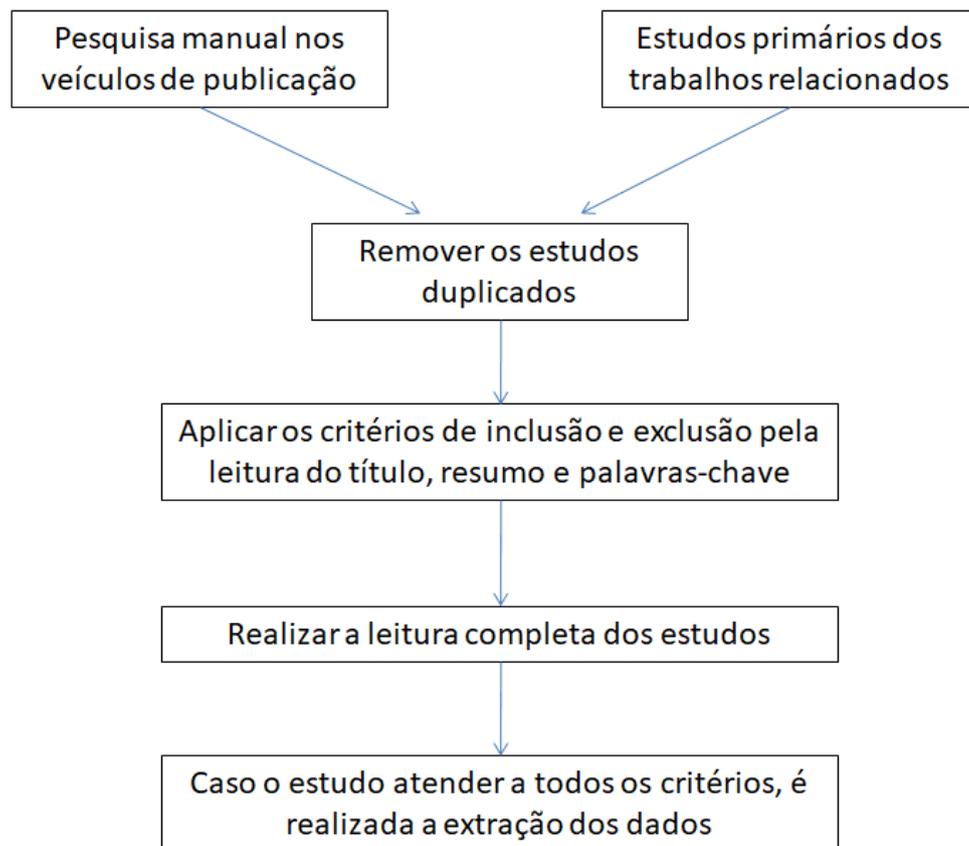
Em nosso trabalho, o processo de triagem foi realizado inicialmente por um pesquisador, efetuando a leitura integral de cada estudo primário selecionado, após aplicação dos critérios de seleção descritos anteriormente. Este pesquisador preencheu o formulário de extração de dados, seguindo o modelo no Anexo I deste trabalho.

Um segundo pesquisador, de forma aleatória, realizou a leitura de 20% dos estudos primários resultantes do critério de seleção previamente aplicado. Caso houvesse diferenças nos resultados obtidos pelos dois pesquisadores em mais de um terço das amostras selecionadas, como, por exemplo, as fases do processo de experimentação que são cobertas pelas ferramentas, o primeiro pesquisador deveria realizar uma avaliação de todos os estudos primários, e assim este processo se repetiria até que não houvessem diferenças na avaliação entre os dois pesquisadores.

Os critérios de inclusão foram aplicados no título do estudo primário, no resumo e nas palavras chaves relacionadas, sendo que todos os critérios devem ser satisfeitos. A exclusão de estudos primários foi realizada apenas após a leitura completa do estudo, quando verificado que o referido estudo realmente não se tratava sobre uma ferramenta de apoio a experimentação, ou quando se aplicava a algum outro critério de exclusão.

O processo de triagem foi realizado conforme o fluxo da Figura 3.2.

Figura 3.2: Processo de triagem de estudos primários



## 3.5 Extração de Dados

A extração de dados é o momento em que as informações referentes às ferramentas são obtidas e processadas, através de uma leitura completa dos estudos primários selecionados realizada pelos pesquisadores, e atendendo às necessidades levantadas pelas questões de pesquisa.

Para realizar a extração dos dados dos estudos primários, preenchemos formulários com os campos necessários para responder as questões de pesquisa, sendo um para cada estudo primário identificado. O formulário utilizado se encontra no Anexo I deste trabalho, e os formulários preenchidos estão no Anexo II.

O formulário está dividido em seções, sendo os cinco primeiros campos referentes à seção de metadados, como título, autores, veículo e ano da publicação. A segunda seção apresenta o quadro com as fases de um projeto de experimentos, e colunas respectivas às questões de pesquisa QP2, QP2.1, QP2.2 e QP2.3, sobre fases suportadas, nível de abstração da linguagem, nível de automação e confiabilidade oferecida.

A última seção do formulário de extração de dados trata sobre as técnicas utilizadas pela ferramenta, como seu funcionamento e principais características, referentes à questão de pesquisa QP1. Também está contemplada nesta seção a questão de pesquisa QP3, sobre o tipo de experimento apoiado pela ferramenta, sendo para experimentos baseados em tecnologia ou não.

Caso o estudo primário não cite em seu conteúdo os dados necessários para determinados campos do formulário de extração, estes foram preenchidos com um traço horizontal " - ", demonstrando que nenhuma informação foi encontrada na extração para aquele campo.

## 3.6 Análise e Síntese dos Dados

Para a análise e síntese de dados, realizamos análise qualitativa e quantitativa, gerando gráficos e tabelas que demonstrassem, de forma mais clara e objetiva, a relação de ferramentas e o suporte oferecido por elas.

Realizamos diferentes cruzamentos das informações levantadas pela extração, e efetuamos as análises qualitativa e quantitativa. Pela análise qualitativa foi possível identificar as similaridades e diferenças das ferramentas de apoio à experimentação, quais os problemas mais comuns encontrado por elas, fases dos projetos de experimentos mais suportadas, suas formas de automação e níveis de linguagens de especificação.

Através da catalogação dos dados em tabelas e gráficos, realizamos a análise quantitativa, que ajudou a gerar informações sobre o número de ferramentas que suportam

experimentos baseados em tecnologia, assim como realizar uma análise combinada com os dados qualitativos para obter informações relacionadas a outras questões de pesquisa.

No Capítulo 4, expomos os resultados desse trabalho e a análise dos dados obtidos, individualmente para cada questão de pesquisa, como também o cruzamento desses dados.

# Capítulo 4

## Resultados e Análise

Neste capítulo reportamos todos os resultados encontrados sobre a aplicação do protocolo de pesquisa, as ferramentas identificadas pela extração de dados, a análise geral do mapeamento e o cruzamento das informações obtidas com relação às questões de pesquisa. As Seções 4.1 e 4.2 sintetizam os dados demográficos do estudo. Já a Seção 4.3 traz os dados de cada questão de pesquisa e a Seção 4.4 apresenta uma análise geral sobre as ferramentas.

### 4.1 Resultados do Protocolo de Pesquisa

A pesquisa do nosso mapeamento sistemático foi realizada em dez veículos de publicações de artigos acadêmicos, listados na Tabela 3.2. Como nosso estudo utilizou como base a revisão sistemática realizada por Freire et al. (2013), aplicamos os nossos critérios de inclusão e exclusão nos estudos primários identificados em seu trabalho. Os estudos que tiveram os nossos critérios satisfeitos, selecionamos para realizar a extração de dados e responder as questões de pesquisa propostas em nosso mapeamento sistemático.

Dessa forma os veículos Journal of Empirical Software Engineering, Journal of Systems and Software, ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM), Experimental Software Engineering Latin American Workshop (ESELAW), e o International Conference on Software Engineering (ICSE) para o período de janeiro de 2002 à dezembro de 2011, já foram pesquisados por Freire et al. (2013) e não realizamos a pesquisa manual nesta combinação novamente. Para esses veículos apenas realizamos a pesquisa manual para um novo período, posterior ao utilizado naquela revisão sistemática, sendo compreendido então o novo período de janeiro de 2012 à janeiro de 2018.

No Capítulo 5, na Seção 5.2 sobre Trabalhos Relacionados, detalhamos o diferencial das nossas questões de pesquisa, e do nosso trabalho como um todo, quanto ao de Freire et al. (2013).

Os veículos IEEE Transactions on Software Engineering (TSE), ACM Transactions on Software Engineering and Methodology (TOSEM), Journal on Software and Systems Modeling (SoSyM), International Conference on Automated Software Engineering (ASE) e European Software Engineering Conference and the Symposium on the Foundations of Software Engineering (ESEC/FSE) não foram compreendidos no trabalho de Freire et al. (2013), e assim realizamos a pesquisa manual destes no período de janeiro de 2002 à janeiro de 2018, aplicando os critérios definidos em nosso protocolo de pesquisa descrito no Capítulo 3.

A Tabela 4.1 apresenta respectivamente os veículos de publicação, os períodos compreendidos em cada um e a quantidade de estudos que foram pesquisados manualmente.

Tabela 4.1: Veículos de publicação, período pesquisado e quantidade de estudos

<b>Veículo de Publicação</b>	<b>Período Pesquisado</b>	<b>Quantidade de Estudos</b>
IEEE Transactions on Software Engineering (TSE)	01/2002 à 01/2018	1648
ACM Transactions on Software Engineering and Methodology (TOSEM)	01/2002 à 01/2018	387
Journal on Software and Systems Modeling (SoSyM)	01/2002 à 01/2018	767
International Conference on Automated Software Engineering (ASE)	01/2002 à 01/2018	1078
European Software Engineering Conference and the Symposium on the Foundations of Software Engineering (ESEC/FSE)	01/2002 à 01/2018	846
Journal of Empirical Software Engineering (ESEJ)	01/2012 à 01/2018	451
Journal of Systems and Software (JSS)	01/2012 à 01/2018	1460
ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)	01/2012 à 01/2018	223
Experimental Software Engineering Latin American Workshop (ESELAW)	01/2012 à 01/2018	98
International Conference on Software Engineering (ICSE)	01/2012 à 01/2018	1290

Pela grande quantidade de material encontrado, a fase de pesquisa ocupou a maior parte do tempo necessário para a elaboração dessa dissertação.

## 4.2 Estudos Primários Seleccionados e Ferramentas Identificadas

Executando o protocolo de pesquisa, realizamos a extração de dados sobre ferramentas de apoio a experimentos nos mesmos estudos primários utilizados no trabalho de Freire et al. (2013), assim como em novos estudos encontrados.

Dentre os quatorze estudos seleccionados, identificamos sete ferramentas de software que se encaixam em nosso mapeamento sistemático. Dentre os quinze estudos citados por Freire et al. (2013), dois se referem apenas a Frameworks para realização de experimentos, sem utilização específica de alguma ferramenta, e portanto não foram incluídos em nosso trabalho. Nesta pesquisa não realizamos a contagem dos estudos que foram lidos por completo e excluídos quando aplicados os critérios de exclusão. Apenas contamos o total de estudos encontrados pela busca manual, e os estudos primários que foram seleccionados após aplicar todos os critérios definidos no protocolo de pesquisa.

A Tabela 4.2 detalha os estudos primários que utilizamos, assim como as ferramentas que foram propostas. Os dois últimos itens da tabela são os novos estudos que foram encontrados por este mapeamento sistemático, e os anteriores são os que foram seleccionados do trabalho de Freire et al. (2013). A Figura 4.1 apresenta a distribuição temporal desses estudos ao longo dos anos.

Tabela 4.2: Estudos primários e ferramentas

<b>ID</b>	<b>Autores</b>	<b>Estudo Primário</b>	<b>Nome da Ferramenta</b>
ES1	Sjoberg et al. (2002)	Conducting realistic experiments in software engineering	Simula Experiment Support Environment (SESE)
ES2	Arisholm et al. (2002a)	A Web-based Support Environment for Software Engineering Experiments	Simula Experiment Support Environment (SESE)
ES3	Arisholm et al. (2002b)	SESE – an Experiment Support Environment for Evaluating Software Engineering Technologies	Simula Experiment Support Environment (SESE)
ES4	Neto et al. (2004)	Infrastructure for SE Experiments Definition and Planning	experimental Software Engineering Environment (eSEE)

ES5	Mian et al. (2004a)	eSEE: a Computerized Infrastructure for Experimental Software Engineering	experimental Software Engineering Environment (eSEE)
ES6	Mian et al. (2005)	A computerized infrastructure for supporting experimentation in software engineering	experimental Software Engineering Environment (eSEE)
ES7	Chapetta et al. (2005)	Supporting Meta-Description Activities in Experimental Software Engineering Environments	experimental Software Engineering Environment (eSEE)
ES8	Travassos et al. (2008)	An environment to support large scale experimentation in software engineering	experimental Software Engineering Environment (eSEE)
ES9	Mian et al. (2004b)	Towards a Computerized Infrastructure for Managing Experimental Software Engineering Knowledge	experimental Software Engineering Environment (eSEE)
ES10	Stolee e Elbaum (2010)	Exploring the use of crowdsourcing to support empirical studies in software engineering	Mechanical Turk
ES11	Hochstein et al. (2008)	An Environment for Conducting Families of Software Engineering Experiments	Experiment Manager framework
ES12	Torii et al. (1999)	Ginger2: An Environment for Computer-Aided Empirical Software Engineering	CAESE Framework and Ginger2
ES13	<b>Hauck et al. (2014)</b>	<b>Deriving performance-relevant infrastructure properties through model-based experiments with Ginpex</b>	<b>Goal-oriented INfrastructure Performance EXperiments (Ginpex)</b>
ES14	<b>Wang et al. (2005)</b>	<b>Automating Experimentation on Distributed Testbeds</b>	<b>Weevil</b>

Para este mapeamento sistemático, levamos em conta a identificação das capacidades das ferramentas apenas pela leitura dos estudos primários, e não pelo uso das ferramentas em si. Por este motivo, quando a leitura do estudo primário que apresenta originalmente a ferramenta não tiver muitos detalhes de seu funcionamento, utilizamos outros estudos que fazem referência à mesma ferramenta para melhorar a compreensão de suas funcionalidades.

A Tabela 4.3 apresenta respectivamente os veículos de publicação e a quantidade de estudos primários que foram selecionados. Os veículos de publicação apresentados são

os que foram utilizados na busca manual, incluindo os estudos selecionados a partir do trabalho de Freire et al. (2013).

Tabela 4.3: Veículos de publicação e quantidade de estudos primários selecionados

<b>Veículo de Publicação</b>	<b>Quantidade de Estudos</b>
IEEE Transactions on Software Engineering (TSE)	1
ACM Transactions on Software Engineering and Methodology (TOSEM)	0
Journal on Software and Systems Modeling (SoSyM)	1
International Conference on Automated Software Engineering (ASE)	1
European Software Engineering Conference and the Symposium on the Foundations of Software Engineering (ESEC/FSE)	0
Journal of Empirical Software Engineering (ESEJ)	0
Journal of Systems and Software (JSS)	0
ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)	1
Experimental Software Engineering Latin American Workshop (ESELAW)	4
International Symposium on Empirical Software Engineering (ISESE)	1
Nordic Workshop on Programming and Software Development Tools and Techniques (NWPER)	1
Nordic Journal of Computing (NJC)	1
IEEE International Conference on Engineering of Complex Computer Systems (ICECCS)	1
Advances in Computers	1
Jornadas Iberoamericanas em Ingeniería del Software e Ingeniería del Conocimiento (JIISIC)	1
International Conference on Software Engineering (ICSE)	0

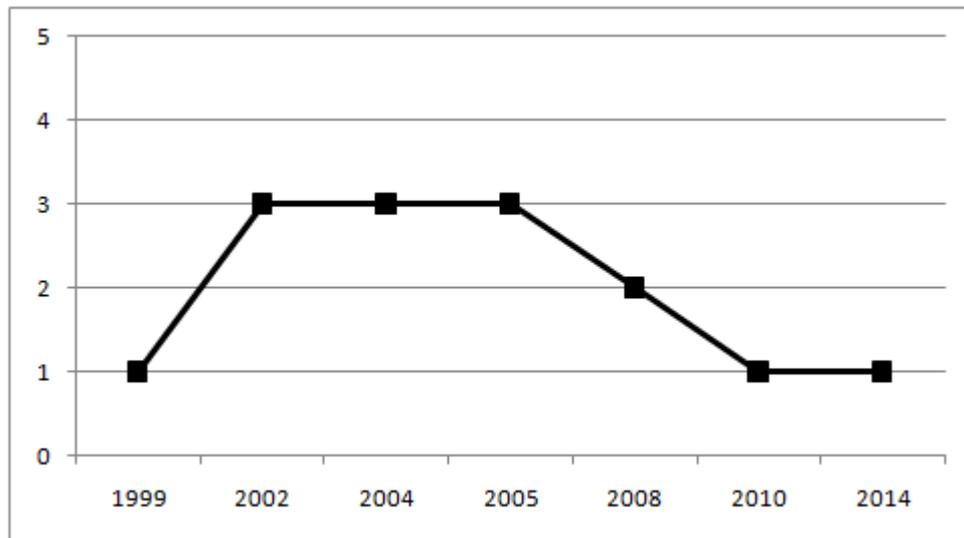
A Tabela 4.4 apresenta o resultado da extração de dados das ferramentas identificadas, e a relação com as questões de pesquisa QP2, QP2.1 e QP2.2, que descrevem quais as fases de um projeto de experimentos são suportadas, qual o nível de abstração da linguagem utilizadas para criação do experimento e o nível de automação oferecido pela ferramenta.

Tabela 4.4: Ferramentas, fases de projeto de experimento suportadas, nível de automação e nível de abstração de linguagem

<b>Ferramenta</b>	<b>Fase de Experimento</b>	<b>Oferece Apoio</b>	<b>Nível de Abstração</b>	<b>Nível de Automação</b>
SESE	Escopo	Sim	Alto	Manual
	Planejamento	Sim	Alto	Manual
	Execução	Sim	Alto	Manual
	Análise e Interpretação	Não	-	-
	Apresentação	Sim	Alto	Semiautomático
eSEE	Escopo	Sim	Alto/Baixo	Semiautomático
	Planejamento	Sim	Alto/Baixo	Semiautomático
	Execução	Sim	Alto	Manual
	Análise e Interpretação	Não	-	-
	Apresentação	Sim	Alto	Semiautomático
Ginger2	Escopo	Não	-	-
	Planejamento	Não	-	-
	Execução	Sim	Alto	Manual
	Análise e Interpretação	Sim	Alto	Semiautomático
	Apresentação	Não	-	-
Experiment Manager Network	Escopo	Não	-	-
	Planejamento	Não	-	-
	Execução	Sim	Alto	Manual
	Análise e Interpretação	Sim	Alto	Semiautomático
	Apresentação	Não	-	-
Mechanical Turk	Escopo	Sim	Alto	Manual
	Planejamento	Sim	Alto/Baixo	Manual
	Execução	Sim	Alto	Manual
	Análise e Interpretação	Não	-	-
	Apresentação	Não	-	-
Ginpex	Escopo	Sim	Alto	Manual/Autom.
	Planejamento	Sim	Alto/Baixo	Manual/Autom.
	Execução	Sim	Alto	Semiaut./Autom.
	Análise e Interpretação	Sim	Alto/Baixo	Manual/Autom.
	Apresentação	Não	-	-
Weevil	Escopo	Não	-	-
	Planejamento	Sim	Alto	Semiautomático
	Execução	Sim	Alto	Automático
	Análise e Interpretação	Sim	Alto	Semiautomático
	Apresentação	Sim	Alto	Semiautomático

Na Figura 4.1 verificamos que o maior número de publicações que apresentassem ferramentas de apoio a experimentos em Engenharia de Software ocorreu entre 2002 e 2005, somando um total de nove estudos primários.

Figura 4.1: Número de estudos primários mapeados e ano de publicação



### 4.3 Resultados Sobre as Questões de Pesquisa

Nesta seção, apresentamos os resultados do mapeamento sistemático com base nas questões de pesquisa levantadas na Tabela 3.1. Para cada questão de pesquisa, relatamos os resultados encontrados pelas funcionalidades suportadas nas ferramentas de apoio a experimentos, baseados no detalhamento apresentado pela Tabela 4.4, resultado da extração de dados nos formulários preenchidos, contantes no Anexo II.

#### 4.3.1 Questão de Pesquisa 1 - Quais são as ferramentas de apoio à experimentação em Engenharia de Software e as funcionalidades de experimentação suportadas pelas ferramentas?

O mapeamento sistemático realizado neste trabalho identificou sete ferramentas que se enquadram nos quesitos definidos no protocolo de pesquisa, descrito no Capítulo 3. Relacionamos na Tabela 4.5 a lista das ferramentas mapeadas, como resultado sobre a questão de pesquisa QP1 do nosso trabalho.

Algumas ferramentas são mais voltadas a verificar parâmetros não funcionais, como dados do sistema operacional e uso da máquina durante um processamento, com objetivo

de executar experimentos relacionados à ambientes computacionais. Outras ferramentas são mais voltadas a verificar a utilização do sistema objeto do experimento por usuários, como, por exemplo, avaliar o funcionamento de um sistema quando executado por usuários com perfis diferentes.

A ferramenta SESE tem como principal característica a criação de formulários Web com questões referentes ao experimento desejado. Participantes preenchem a sequencia de questões e submetem suas respostas. A ferramenta gerencia os participantes, monitorando a execução e o tempo utilizado no experimento.

Sobre a ferramenta eSEE, sua principal característica é a infraestrutura para um ambiente experimental em Engenharia de Software, seguindo um modelo para empacotamento de experimentos, que define uma taxonomia de documento para representar a informação necessária a execução de experimentos. Essa infraestrutura é definida para suportar processos de experimentação em larga escala, fornecendo recursos para que experimentadores em Engenharia de Software, em locais diferentes, possam realizar experimentos através de uma rede de computadores.

Experiment Manager é uma infraestrutura em rede, formada por um conjunto de ferramentas, que simplifica o processo de coletar, gerenciar e normalizar dados de experimentos em engenharia de software, executados em salas de aula, e minimiza a mudança comportamental natural dos indivíduos relacionados ao experimento.

O framework CAESE, utilizado pela ferramenta Ginger2, cobre todas as fases de um projeto de experimentação, porém a ferramenta implementa apenas parte do framework, oferecendo apoio para as fases de Execução e Análise e Interpretação.

A ferramenta Mechanical Turk é um software de crowdsourcing que foi utilizado no estudo primário de Stolee e Elbaum (2010) para obter resultados sobre um experimento envolvendo muitos indivíduos, usuários do software. Com ela é possível determinar o tipo e a quantidade de indivíduos que aplicarão tratamentos ao objeto do experimento. Mechanical Turk é um exemplo de ferramenta que não foi desenvolvida propriamente para realizar experimentos, mas que pode ser utilizada para este fim conforme apresentado no seu estudo primário.

Ginpex é direcionada a experimentos baseados em tecnologia, que verificam parâmetros como nível de performance de um sistema, tempo de processamento, uso de recursos computacionais, dentre outros.

Por último, a ferramenta Weevil, que tem como principal característica a execução de experimentos em sistemas distribuídos, usando preferencialmente redes computacionais chamadas Testbeds, próprias para avaliação de softwares que necessitam controlar acessos remotos e troca de dados entre diversas máquinas clientes.

Tabela 4.5: Resultado da questão de pesquisa 1 - Lista de ferramentas

#	Nome da Ferramenta
01	Web-based Simula Experiment Support Environment (SESE)
02	experimental Software Engineering Environment (eSEE)
03	Mechanical Turk
04	Experiment Manager framework
05	CAESE Framework and Ginger2
06	Goal-oriented INfrastructure Performance EXperiments (Ginpex)
07	Weevil

### 4.3.2 Questão de Pesquisa 2 - Quais são as fases de um projeto de experimentação cobertas pela ferramenta?

Nas cinco fases de um projeto de experimentos em engenharia de software, definidos por Wohlin et al. (2012), identificamos pelo menos uma ferramenta que as suportassem, sendo elas Escopo, Planejamento, Execução, Análise e Interpretação e, por fim, Apresentação e Empacotamento. Porém também identificamos que nenhuma ferramenta suporta todas as fases de um projeto de experimentação. A Figura 4.2 apresenta a quantidade de ferramentas em que cada fase é suportada.

A fase de Escopo é suportada por quatro das ferramentas levantadas pelo mapeamento sistemático, sendo elas SESE, eSEE, Mechanical Turk e Ginpex. Esta também é uma fase em que há maior necessidade de intervenção humana, pois trata da elaboração do que o projeto de experimento irá realizar.

O Planejamento é suportado por cinco ferramentas, sendo a segunda fase com maior apoio oferecido por elas. As ferramentas que suportam são: SESE, eSEE, Mechanical Turk, Ginpex e Weevil

A fase de Execução é suportada por todas as ferramentas, portanto é possível interpretar essa informação como sendo esta fase a mais crítica para uma ferramenta de apoio a experimentação, ou porque as outras fases são mais difíceis de serem implementadas, ou ainda porque as ferramentas estão em estágio de evolução.

Verificamos que há um número maior de ferramentas que suportam as fases de escopo, planejamento e execução. Podemos interpretar isso como, provavelmente, ocorrendo devido ao impacto que essas fases apresentam na especificação de um projeto de experimentação.

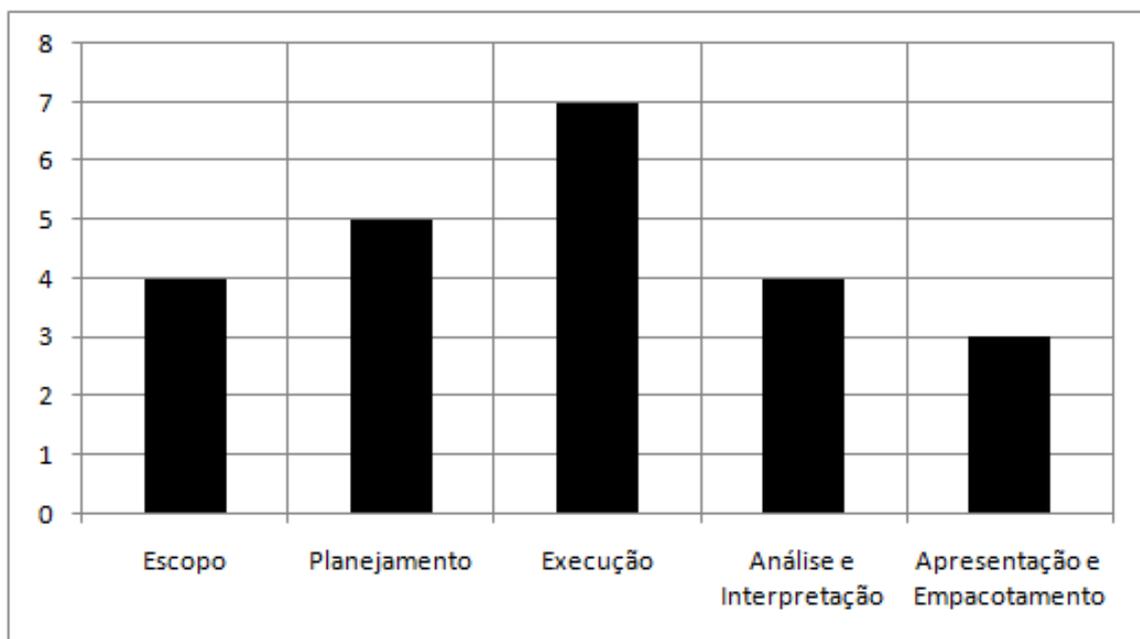
Para Análise e Interpretação, quatro ferramentas suportam e dão apoio a esta fase, são elas Ginger2, Experiment Manager Network, Ginpex e Weevil. Podemos verificar pelo mapeamento que as ferramentas que implementam funcionalidades para análise são as que oferecem algum nível de automação para ela. Podemos concluir que quando ferramentas

trabalham os dados obtidos durante a execução do experimento, alguma inteligência deve ser adicionada para realizar esse procedimento. As ferramentas que não oferecem suporte a esta fase determinam que o próprio experimentador deve trabalhar os dados para extrair seus resultados para análise.

A fase de Apresentação e Empacotamento também foi uma das com menor implementação pelas ferramentas. Apenas três a suportam, sendo elas SESE, eSEE e Weevil. Isso pode ter sido ocasionado pela necessidade maior de interferência humana no processo de elaborar a apresentação dos resultados, mas a parte de empacotamento, que auxilia na reprodução do experimento, seria algo que a própria ferramenta poderia fornecer.

Por fim, as quatro ferramentas que suportam mais fases de projetos de experimentos são: SESE, eSEE, Ginpex e Weevil. Estas apoiam o pesquisador em Engenharia de Software em quatro das cinco fases.

Figura 4.2: Resultado da questão de pesquisa 2 - Quantidade de ferramentas e fases suportadas



### 4.3.3 Questão de Pesquisa 2.1 - Qual o nível de abstração da linguagem de especificação da ferramenta?

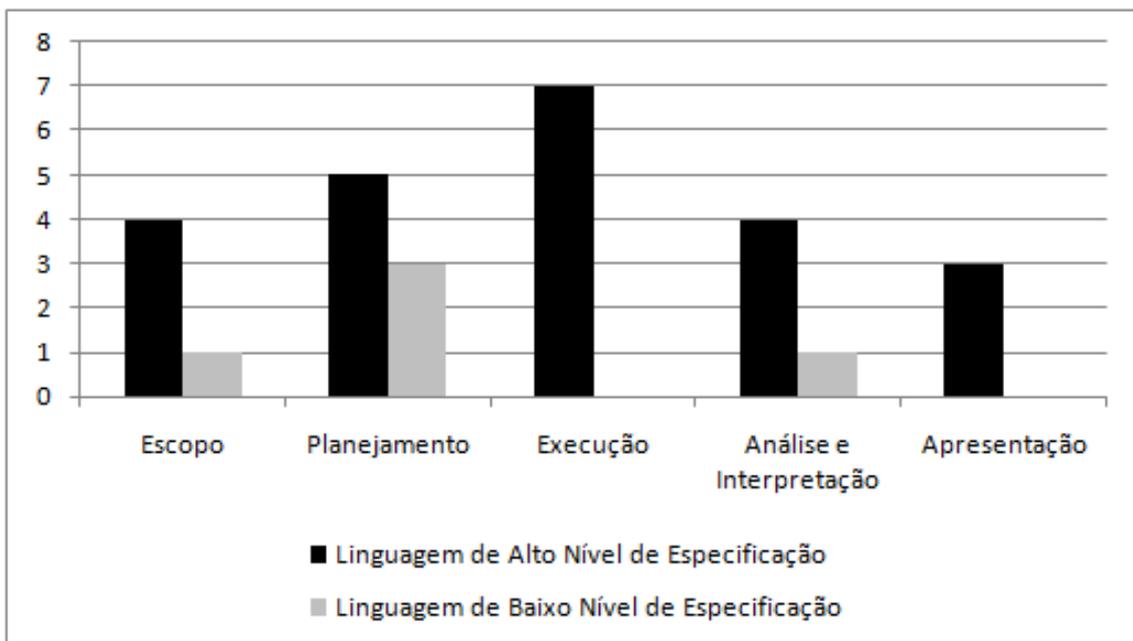
Dentre as ferramentas identificadas em nosso mapeamento sistemático, as que oferecem suporte a linguagem de alto nível em alguma das fases de um processo de experimento, em geral, utilizam elaboração por modelos. Um exemplo é a ferramenta Weevil, que, através de uma abordagem baseada em modelos, define o sistema em experimentação, qual a

rede computacional que será utilizada, e os comportamentos dos clientes que conduzem o experimento. Isso ocorre principalmente na fase de Planejamento.

Já as fases que suportam linguagem de baixo nível, como, por exemplo, a ferramenta Ginpex, são ou programadas pelo experimentador desenvolvedor ou fornecem templates pré desenvolvidos que podem ser customizados.

A Figura 4.3 apresenta a quantidade de ferramentas em que cada fase é suportada, divididas pelo nível de abstração da linguagem de especificação. Na barra esquerda de cada fase, demonstra a quantidade de ferramentas que suportam linguagem de alto nível, e na barra direita as que suportam linguagem de baixo nível.

Figura 4.3: Resultado da questão de pesquisa 2.1 - Quantidade de ferramentas por nível de abstração da linguagem de especificação para cada fase suportada



Quando é observada a fase de Escopo, o nível de abstração da linguagem para sua especificação é alto nível, pois basicamente trata da especificação do que o experimento pretende atender, e isso pode ser obtido através da escrita com linguagem natural.

A fase de Planejamento é a que mais oferece suporte tanto a linguagens de baixo, como também para as de alto nível, a disposição do experimentador escolher qual atende melhor sua necessidade. Isso se deve pela disponibilização de elaboração desta fase no experimento por modelos visuais ou templates, como também pela codificação em linguagens de programação, caso seja necessária alguma customização maior do que deverá ser executado.

Em todas as ferramentas pode-se notar que, pela necessidade de controle e gerência da execução do experimento ser administrada pelas próprias ferramentas, a fase de execução

oferece somente suporte a linguagens de alto nível, sem muita liberdade para que o experimentador realize customizações. Todas as ferramentas oferecem suporte a linguagens de alto nível em alguma fase, mas as ferramentas eSEE, Mechanical Turk e Ginpex são as que oferecem suporte a linguagem de baixo nível, em pelo menos uma das cinco fases do experimento.

Para exemplo de uma ferramenta que implementa tanto a linguagem de especificação de baixo nível como a de alto nível em uma fase, citamos a ferramenta Ginpex, que permite ao engenheiro de software configurar em códigos-fonte os tratamentos a que os sujeitos do experimento serão submetidos, mas dessa forma a fase de Análise não será especificada pelo Ginpex para que seja interpretada de forma totalmente automatizada. Caso sejam utilizados modelos, a fase de Análise, interpreta automaticamente os resultados obtidos na execução dos experimentos.

#### **4.3.4 Questão de Pesquisa 2.2 - Qual é o nível de automação suportado pela ferramenta?**

A Figura 4.4 apresenta a quantidade de ferramentas em que cada fase é suportada, divididas pelo nível de automação. Na barra esquerda de cada fase, demonstra a quantidade de ferramentas que oferecem suporte manual, na barra do meio as que oferecem suporte semiautomático, e na barra direita as que oferecem suporte automático.

Neste estudo, definimos como suporte semiautomático as ferramentas que utilizam parcialmente a entrada realizada de forma manual pelo experimentador para, então, processar e gerar algum conteúdo para uma ou mais fases do projeto de experimentação. O suporte totalmente manual é identificado quando a ferramenta apenas registra a entrada do usuário para aquela fase do processo, mas não realiza nenhuma operação para gerar mais conteúdo. Por fim, o suporte totalmente automático, definimos como a ferramenta gerando conteúdo, para uma ou mais fases do processo de experimento, baseado em um simples passo de seleção ou escolha em utilizar um template pré-definido na ferramenta para aplicar em seu experimento, pelo pesquisador em Engenharia de Software.

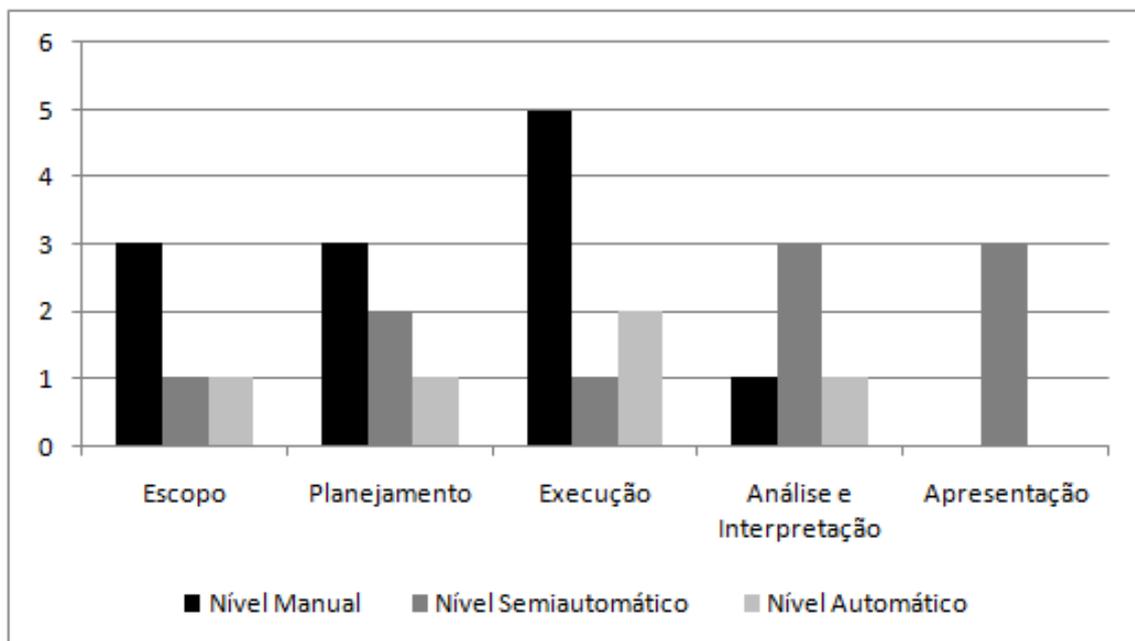
No mapeamento, foi possível identificar que as ferramentas com grau Automático, que suportam pelo menos uma das fases do processo de experimento, são as duas que se propõem a apoiar experimentos baseados em tecnologia, a Ginpex e a Weevil.

Já o número de ferramentas que suportam fases com nível de automação semiautomático é maior, seis no total, sendo elas SESE, eSEE, Ginger2, Experiment Manager Network, Ginpex e Weevil. Essas oferecem apoio parcial para elaboração manual pelo experimentador, e alguma parte é automatizada pela ferramenta.

Analisando os dados obtidos mapeamento sistemático realizado, podemos observar que a maioria das fases compreendidas pelas ferramentas de apoio a experimentos oferecem suporte manual, necessitando assim uma maior intervenção humana para elaboração, execução e análise dos dados de um experimento. Destas, seis ferramentas oferecem suporte manual a alguma fase do processo, sendo elas SESE, eSEE, Ginger2, Experiment Manager Network, Mechanical Turk e Ginpex.

Como exemplo de funcionalidades oferecidas pelas ferramentas para níveis de automação, citamos a Ginpex, que, através de modelos pré-configurados de experimentos, possibilita desde a geração automática de definição de escopo como seu planejamento, execução e análise dos resultados obtidos. Dessa forma o engenheiro de software pode apenas selecionar qual tipo de informação será verificada (tempo de processamento, uso de memória, etc), definir qual sistema em que será executado o experimento e o Ginpex realizará o procedimento.

Figura 4.4: Resultado da questão de pesquisa 2.2 - Quantidade de ferramentas por nível de automação para cada fase suportada



#### 4.3.5 Questão de Pesquisa 2.3 - Qual é a confiabilidade fornecida pela ferramenta?

Para esta pesquisa, a definição de confiabilidade que utilizamos se trata de uma garantia de rastreabilidade entre os resultados da análise do experimento de acordo com a sua especificação. Não tratamos da confiabilidade no sentido de robustez da ferramenta em

caso de falha de software, como, por exemplo, a recuperação do experimento quando o sistema é interrompido.

Com este mapeamento sistemático, identificamos uma grave falta de evidências na validação de resultados em relação à especificação de um experimento. Embora as ferramentas mapeadas auxiliem na condução de experimentos controlados, nenhuma delas oferece garantias da validade de resultados para experimentos. A falta de provas de validade adequadas da ferramenta pode levar a resultados inválidos.

Conforme Juristo e Moreno (2013) e Wohlin et al. (2012), a validade e a replicabilidade devem ser abordadas desde as fases iniciais, para que possam ser alcançadas. Easterbrook et al. (2008) classifica a validade em quatro tipos: validade de construto, que determina se as construções teóricas são corretamente interpretadas e medidas; validade interna, determinando se o planejado e os resultados realmente estão alinhados com os dados obtidos; validade externa, que é relacionada à generalização dos resultados; e a confiabilidade, relacionada ao viés na pesquisa.

Ferramentas de apoio a experimentos em Engenharia de Software que apresentam confiabilidade, devem considerar na fase de planejamento a validade dos resultados. Na fase de análise e interpretação, devem avaliar os resultados de execução que realmente correspondem às hipóteses definidas na especificação do experimento, usando uma função de análise adequada e os parâmetros corretos. E na fase de apresentação e empacotamento, todos os scripts e dados brutos devem ser apresentados, para que outros pesquisadores que replicarem o experimento produzam os mesmos resultados da execução original.

Os resultados do processo geral de experimentação devem ser consistentes com a especificação do experimento. Nenhuma das ferramentas identificadas, no mapeamento sistemático, fornece evidências da validade dos resultados fornecidos por suas abordagens. Mesmo que existam evidências empíricas, ainda há falta de garantias da validade dos resultados em relação à especificação do experimento.

Como os estudos primários mapeados não citavam claramente sobre confiabilidade nas fases suportadas pelos sistemas de apoio a experimentos, nosso mapeamento sistemático identificou isso como uma sugestão para trabalhos futuros. Novos estudos podem ser realizados sobre a capacidade de confiabilidade nessas ferramentas.

#### **4.3.6 Questão de Pesquisa 3 - Que tipo de experimento é apoiado pela ferramenta?**

Um dos focos deste estudo é identificar quais ferramentas oferecem apoio a experimentos baseados em tecnologia. Conforme a definição de Wohlin et al. (2012), nesse tipo de experimento uma tecnologia é o indivíduo que executa os tratamentos nos objetos, e dessa

forma o controle sobre a execução é maior, pois a tecnologia tende a ser determinística. A Figura 4.5 apresenta a quantidade de ferramentas que oferecem suporte ou não a experimentos baseados em tecnologia.

Um exemplo desse tipo de experimento seria a execução de duas ferramentas que geram casos de teste, em um mesmo programa, e assim comparar qual tem melhor desempenho. Experimentos que não são baseados em tecnologia dependem de humanos para sua execução, o que fornece menor controle sobre o experimento, pois cada ser humano tem capacidades distintas de acordo com seu aprendizado, assim ocorrendo um viés maior do que se executado por programas.

Dentre as sete ferramentas mapeadas em nossa pesquisa, cinco delas não apresentaram, em seus estudos primários, que efetivamente implementam capacidade de execução do experimento de forma semiautomática ou automática. Estas ferramentas são: SESE, eSEE, Ginger2, Experiment Manager Network e Mechanical Turk. Nas implementações de experimentos propostos nestes estudos primários, as ferramentas descrevem explicitamente a execução por um humano, como, por exemplo, no preenchimento de questionários web, ou na implementação de códigos-fonte entre desenvolvedores de diversos níveis de conhecimento.

A ferramenta eSEE, em sua arquitetura, prevê o acoplamento de um módulo para execução automática de experimentos *in silico*, mas no estudo primário de Travassos et al. (2008) descreve que este módulo do eSEE ainda está em desenvolvimento.

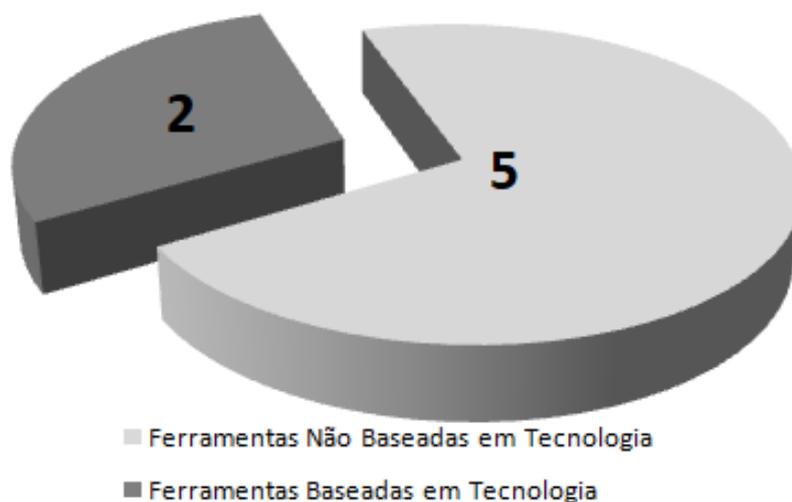
Apenas duas das ferramentas mapeadas oferecem suporte a experimentos baseados em tecnologia, a Ginpex e a Weevil, que são aquelas que se propõe a avaliar parâmetros não funcionais, assim como o resultado da questão de pesquisa QP2. Essas ferramentas permitem a execução do experimento sem intervenção humana.

Uma delas, a Weevil, se não oferecesse suporte baseado em tecnologia, demandaria que centenas ou milhares de usuários executassem ações simultaneamente em um sistema a partir de computadores distintos, para que conseguisse obter os mesmos resultados. A outra, Ginpex, executa experimentos voltados a processamento, uso de memória, dentre outros parâmetros computacionais.

## 4.4 Análise Geral das Ferramentas

Com base na extração de dados realizada, geramos dois gráficos, um com a representação de ferramentas que oferecem apoio a experimentos baseados em tecnologia (QP3) e outro com as ferramentas de experimentos não baseados em tecnologia. Geramos esses gráficos para correlacionar os resultados das questões de pesquisa, buscando padrões e oportunidades de melhoria. Ambos os gráficos das Figuras 4.6 e 4.7 demonstram os resultados encontrados

Figura 4.5: Resultado da questão de pesquisa 3 - Quantidade de ferramentas por tipo de experimento suportado



neste mapeamento dividido em três eixos, de acordo com as questões de pesquisa QP2, QP2.1 e QP2.2.

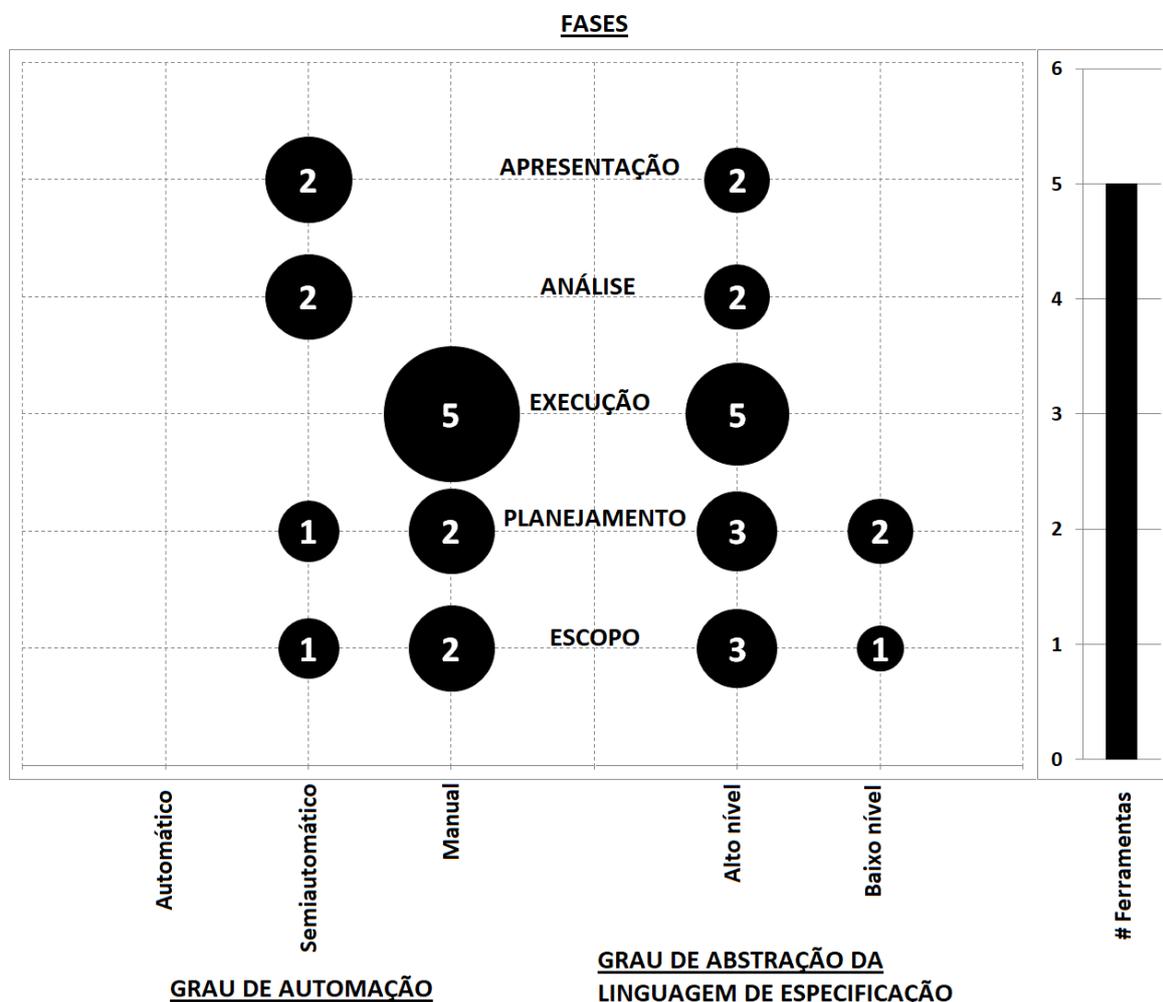
Nas figuras, o eixo vertical representa as cinco fases de um projeto de experimentos que podem ser suportadas pelas ferramentas, sendo elas escopo, planejamento, execução, análise e apresentação. O eixo da horizontal esquerda trata do nível de automação das ferramentas para cada uma das fases do eixo principal, sendo manual, semiautomático e automático. O eixo da horizontal direita apresenta o nível de abstração da linguagem de especificação utilizada em cada ferramenta, também para cada fase do eixo principal deste gráfico, e suas opções são linguagens de alto nível e de baixo nível. Por fim, em ambas as figuras, um gráfico em barra demonstrando a quantidade total das ferramentas está presente ao lado direito.

#### 4.4.1 Ferramentas de Apoio a Experimentos Não Baseados em Tecnologia

Dentre as sete ferramentas identificadas, cinco delas não apoiam experimentos baseados em tecnologia, sendo elas SESE, eSEE, Ginger2, Experiment Manager Network e Mechanical Turk. Verificamos que nenhuma oferece suporte a todas as fases de um projeto de experimentos. Este é o primeiro ponto de falha que identificamos para evolução das ferramentas.

Pelo gráfico demonstrado na Figura 4.6, é possível verificar que não foi encontrada nenhuma ferramenta que ofereça suporte a qualquer fase com grau totalmente automático.

Figura 4.6: Fases suportadas, nível de automação e nível de abstração de linguagem para ferramentas de apoio a experimentos Não Baseados em Tecnologia



Isso tende a ocorrer devido ao tipo de experimento suportado, que não é baseado em tecnologia, logo uma intervenção humana maior é necessária. Esse ponto deixa abertura para um trabalho futuro, que avalie a viabilidade de aumentar o nível de automação em todas as fases dessas ferramentas.

A maior parte dessas ferramentas oferece suporte manual para algumas das fases de um experimento, sendo que todas as cinco tem esse grau de automação para a fase de Execução, e ao mesmo tempo verificamos que todas as ferramentas suportam essa fase com linguagem de especificação de alto nível. Com essa informação podemos concluir novamente que, por se tratarem de ferramentas de apoio a experimentos não baseados em tecnologia, a execução do experimento necessita ser controlada por um ser humano, e assim não tem apoio automatizado para essa tarefa. E, por esse mesmo motivo, não exige a necessidade de especificar as tarefas da fase de Execução com linguagem de baixo nível. Entretanto, um ponto para evolução das ferramentas será aumentar os seus níveis de

automação, de forma que o ser humano consiga executar o experimento com pelo menos algum suporte semiautomatizado, exigindo menor esforço e diminuindo os riscos de erros por falha humana na execução totalmente manual.

Já as fases de Análise e Apresentação são as que tem a menor quantidade de ferramentas que as suportam, apenas duas. Ambas oferecem nível semiautomático para realização das suas tarefas e linguagem de alto nível de abstração. Com isso, podemos inferir que mesmo as ferramentas não suportando o nível totalmente automatizado para realização das tarefas nas fases de Análise e Apresentação, que tornaria o processo mais ágil, ainda torna as atividades do experimentador mais fácil por ter algum grau de automação e com linguagem especificação de alto nível. Demonstra que uma intervenção humana na escolha de opções de análise e apresentação ainda é necessária. A evolução possível para estas fases seria de aumentar o nível de automação.

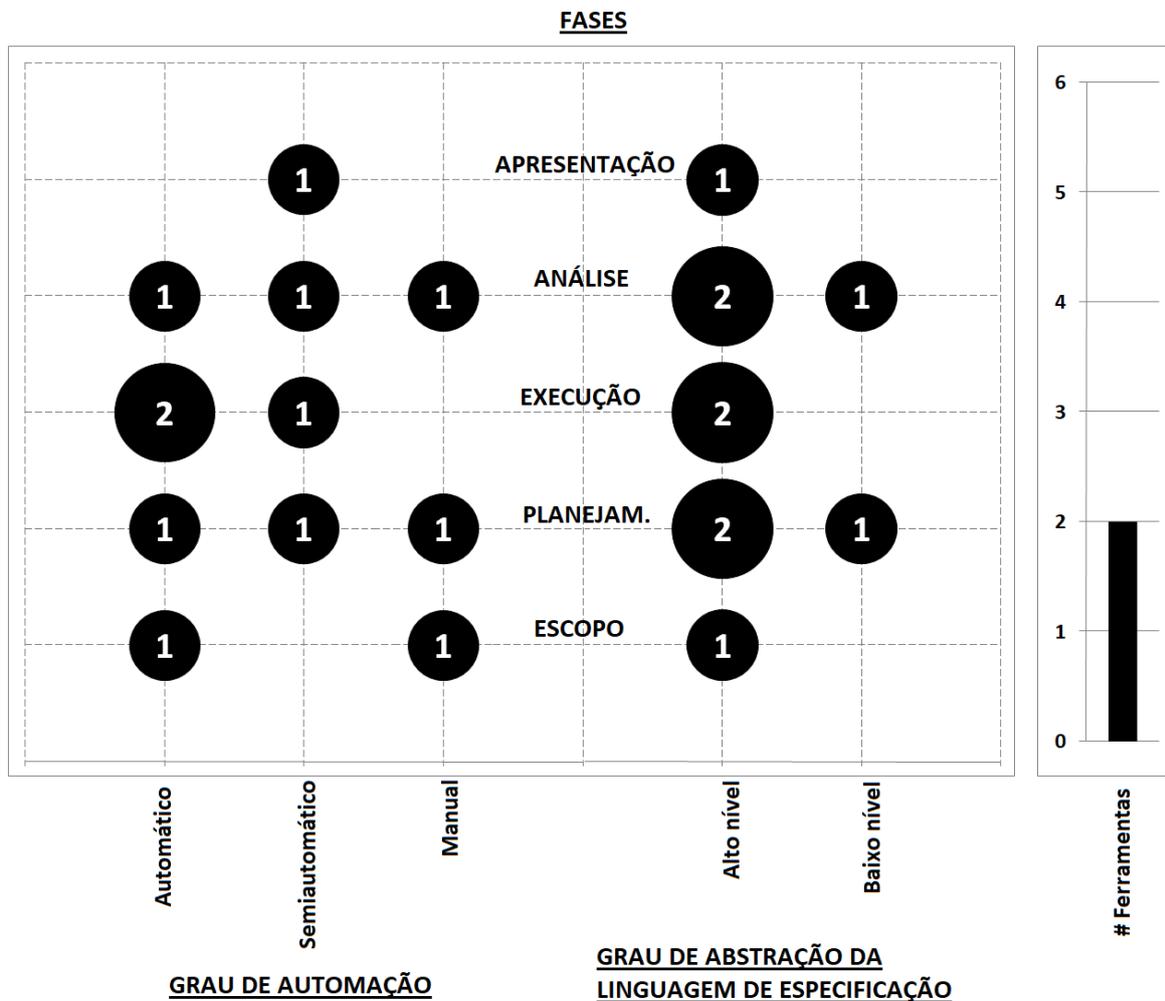
Em ferramentas de apoio a experimentos que não são baseados em tecnologia, é possível identificar que as fases de Escopo, Planejamento e Execução são as que oferecem suporte basicamente de forma manual, com pouco ou nenhum auxílio da ferramenta no que trata de sua automação. Isso se deve pelo fato de que, as fases de Planejamento e Escopo, apresentam um maior grau de definições sobre como o experimento deve ser conduzido, exigindo que as atividades dessas fases possam ser configuradas e customizadas de várias formas para atender aos requisitos dos experimentos não baseados em tecnologia. Também por este motivo verificamos que linguagens de especificação de alto nível prevalecem sobre as de baixo nível, permitindo uma maior configurabilidade na especificação do experimento. Porém nem todas as ferramentas suportam essas fases, deixando aqui um destaque para trabalhos futuros.

A maioria das ferramentas oferece suporte a linguagem de alto nível em pelo menos alguma fase, sendo que todas as cinco compreendem a fase de Execução nesta forma. As fases de Apresentação e de Análise também contemplam apenas linguagens de alto nível, e apenas as fases de Escopo e de Planejamento tem a opção de linguagens de baixo nível. Podemos concluir que, para experimentos que não são baseados em tecnologia, há uma maior dificuldade em especificar as tarefas das fases com linguagem de baixo nível. Este é outro destaque para novos estudos serem desenvolvidos, a viabilidade de implementar uma ferramenta de experimentos não baseados em tecnologia que possa ser especificada com linguagem de baixo nível.

## 4.4.2 Ferramentas de Apoio a Experimentos Baseados em Tecnologia

Nosso mapeamento sistemático levantou apenas duas ferramentas de apoio a experimentos baseados em tecnologia, conforme demonstrado na Figura 4.7. As ferramentas são a Ginpex e a Weevil. Esse é o primeiro ponto de falha que identificamos para evolução das ferramentas de apoio a experimentos, a implementação de mais ferramentas voltadas a experimentos baseados em tecnologia. Por esse motivo a análise de pontos de falha e de ênfase mais comuns para as ferramentas desse tipo, descritos nesta seção, são mais limitados, dificultando afirmar um padrão entre as funcionalidades suportadas e possíveis evoluções.

Figura 4.7: Fases suportadas, nível de automação e nível de abstração de linguagem para ferramentas de apoio a experimentos Baseados em Tecnologia



Verificamos que, apesar de apenas duas ferramentas oferecerem apoio a este tipo de experimento, elas oferecem suporte a mais de uma opção, tanto sobre o nível de automação

como para o nível de abstração da linguagem de especificação, em determinadas fases do processo de experimento. Isso ocorre nas fases de planejamento, execução e de análise e interpretação do experimento. Podemos concluir que, por serem ferramentas direcionadas a experimentos baseados em tecnologia, é maior a possibilidade de especificar as tarefas de um projeto de experimentação com nível automático ou manual, mesmo que a linguagem de especificação seja de alto nível. Sugerimos estudos para verificar o motivo, pela menor importância que se dá, para que nenhuma ferramenta ofereça apoio totalmente automático para a fase de Apresentação. Como exemplo, uma forma dessa implementação pode ser a criação automática de containers ou máquinas virtuais, com o ambiente tecnológico do experimento já configurado, para ajudar na replicação do mesmo experimento ou na reprodução de novos experimentos similares por outros pesquisadores. Mesmo que ofereçam suporte de linguagem de alto nível, um nível de automação maior pode ajudar mais ainda o experimentador na especificação deste fase.

Sobre o nível de abstração das linguagens, utilizadas na especificação do experimento pelas ferramentas, o formato mais comum é a linguagem de alto nível, principalmente pela utilização de modelos para construção de experimentos. Poucas fases podem ser adaptadas pelos pesquisadores em Engenharia de Software para uma maior customização, como por código-fonte, por exemplo. Analisamos que as fases de Planejamento e de Análise oferecem diversos níveis de automação como de abstração de linguagem. Isso pode ocorrer pelo fato de que as ferramentas direcionadas a experimentos baseados em tecnologia sejam em maior parte utilizadas por experimentadores com conhecimentos mais técnicos, e assim oferecem opções para especificar essas fases com linguagens de alto e de baixo nível.

Nessas ferramentas é possível verificar que diversas fases tem suporte automatizado ao experimentador, e com linguagens de alto nível. Podemos inferir que isso ocorre pela natureza dos experimentos baseados em tecnologia, onde a intervenção humana é menor. Sugerimos estudos de viabilidade para verificar implementações de linguagens de especificação de baixo nível para as fases de Escopo, Execução e Apresentação, pois nenhuma das ferramentas identificadas as suportam com esse nível.

### **4.4.3 Discussão e Implicações**

O primeiro ponto que exige atenção, com os dados obtidos por este mapeamento sistemático e detalhados na Tabela 4.4, é a baixa quantidade de ferramentas de apoio a experimentos em Engenharia de Software. Com mais de sete mil estudos processados pela busca manual, apenas sete ferramentas foram identificadas. Muitos estudos pesquisados, utilizando nossos critérios de seleção, citam a realização de experimentos para testar alguma hipótese, mas não descrevem a utilização de uma ferramenta de apoio na sua execução. Dessa forma, identificamos essa necessidade comum a várias pesquisas, mas a dificuldade em implementar

uma ferramenta que suporte fases de um projeto de experimentação, em qualquer nível, é um dos pontos principais levantados por nosso estudo.

Analisando todas as ferramentas identificadas pelo mapeamento, verificamos que aquelas que são voltadas a experimentos baseados em tecnologia tem um maior nível de automação nas fases suportadas, como esperado, já que os experimentos que não são baseados em tecnologia, onde pessoas aplicam os tratamentos aos objetos, exigem um grau maior de interação humana. Mas a falta de suporte com maior nível de automação implica que mais esforço é exigido do pesquisador em Engenharia de Software, quando na elaboração do experimento utilizando uma ferramenta de apoio. Assim, é possível prever que a quantidade de tempo necessária para realizar a tarefa de uma fase com nível de automação manual também será maior. Verificada essa falta de suporte em nível de automação em grande parte das ferramentas, é um ponto para pesquisas futuras e novas propostas para evolução dessas ferramentas.

Dentre as sete ferramentas, apenas três delas oferecem suporte a abstração da linguagem de especificação de baixo nível, como, por exemplo, quando a elaboração de uma fase do experimento pode ser feita utilizando uma linguagem de programação. Podemos inferir que isso ocorre pelo fato de que implementar essa opção, numa ferramenta de apoio a experimentos, seja mais complicada que utilizar a linguagem natural ou orientação por modelos. Este é outro ponto que consideramos como um problema comum às ferramentas identificadas pelo mapeamento.

Avaliando os gráficos das Figuras 4.6 e 4.7, identificamos um ponto de falha em comum, tanto em ferramentas para experimentos baseados em tecnologia ou não, que é o nível de abstração da linguagem de especificação para a fase de Execução ser somente de alto nível. Podemos inferir que há dificuldade, ou não foi encontrada necessidade, de implementação de ferramentas que especifiquem as tarefas da fase de Execução utilizando linguagem de baixo nível. Isso também ocorre para a fase de Apresentação e Empacotamento, que além de ter especificação em linguagem de alto nível, somente oferece suporte em nível semiautomático, sugerindo a necessidade de que o experimentador realize uma intervenção parcial para realização das tarefas desta fase. Estudos de viabilidade, voltados a ferramentas com apoio às fases de Execução e de Apresentação e Empacotamento, são necessários para analisar esses pontos de destaque.

## **4.5 Guia Preliminar para Escolha de Ferramenta**

Nesta seção, apresentamos um guia preliminar para facilitar a escolha de uma ferramenta para realização de um processo de experimento. Na Seção 4.5.1 apresentamos uma tabela com perguntas que auxiliam a identificar as necessidades, do pesquisador em Engenharia

de Software, no experimento que deseja conduzir. Também apresentamos duas tabelas com as características suportadas em cada ferramenta, para que, de acordo com as necessidades levantadas anteriormente, seja mais fácil identificar a ferramenta mais apropriada. Na Seção 4.5.2 descrevemos uma análise resumida sobre as vantagens e desvantagens das ferramentas para cada tipo de experimento, detalhando suas limitações, para que o experimentador o utilize como um filtro final deste guia.

Elaboramos a Figura 4.6, que descreve um pequeno fluxo para utilização deste guia preliminar, determinando a sequência que um pesquisador em Engenharia de Software pode seguir para identificar a ferramenta mais adequada para seu experimento. Primeiro, o experimentador determina se o processo de experimentação que deseja conduzir se refere a um experimento baseado em tecnologia ou não. Pelas características distintas desses tipos de experimento, esse é um filtro que reduz bastante a quantidade de ferramentas que se adequam a cada experimento. O segundo passo do fluxo é determinar as fases que o projeto do experimento a ser conduzido deve ser suportado por uma ferramenta. O terceiro passo é identificar qual o nível de abstração da linguagem de especificação mais apropriado para cada fase do projeto de experimentação, assim como o nível de automação desejado. Com o filtro aplicado pelos passos anteriores, a lista de ferramentas que atendem aos requisitos do experimento já estará bem reduzida. O último passo é verificar as vantagens, desvantagens e limitações de cada ferramenta. Com essas definições é possível filtrar as ferramentas que conseguem apoiar a elaboração do experimento para as fases, os níveis de suporte e características exigidas, ou identificar a ferramenta que mais se aproxime das necessidades especificadas.

#### 4.5.1 Identificação das Necessidades do Experimento e Suporte Oferecido

Nesta seção elaboramos o questionário apresentado na Tabela 4.6 para determinar qual o nível do apoio que o pesquisador em Engenharia de Software busca nas ferramentas levantadas em nosso mapeamento sistemático.

Tabela 4.6: Questionário para identificação das necessidades de ferramenta de apoio a experimentos

<b>Pergunta 01 - Em seu experimento, o indivíduo que aplicará os tratamentos ao objeto definido será um computador/sistema/máquina ou um ser humano?</b>	
Detalhamento da pergunta	Esta pergunta define se o experimento será baseado em tecnologia ou não.

Definição da Resposta	Caso o indivíduo seja um ser humano, utilize a Tabela 4.7, de ferramentas não baseadas em tecnologia, no resto deste guia. Caso o indivíduo seja um computador/sistema/máquina, utilize a Tabela 4.8, de ferramentas baseadas em tecnologia.
<b>Pergunta 02 - No experimento, é necessário suporte de uma ferramenta que possa auxiliar na definição de escopo, como por exemplo descrever o objetivo e a motivação do experimento?</b>	
Detalhamento da pergunta	O escopo é definido através de um objeto a ser experimentado, com um determinado propósito ou intenção, buscando algum efeito com foco em qualidade, sob uma perspectiva de ponto de vista para interpretação dos resultados, e o contexto de pessoas ou tecnologias envolvidas.
Definição da Resposta	Caso esse suporte for necessário para a realização do seu processo de experimentação, então escolha uma ferramenta que apoie a fase de Escopo. A seguir, verifique nas Perguntas 7 e 8 quais os níveis de abstração de linguagem e de automação mais adequados ao seu processo.
<b>Pergunta 03 - No experimento, é necessário suporte de uma ferramenta que possa auxiliar no planejamento de como o experimento deve ser conduzido?</b>	
Detalhamento da pergunta	As tarefas de planejamento podem incluir a definição do ambiente em que o experimento será executado, a formulação da hipótese para análise estatística, a seleção das variáveis dependentes e independentes, a seleção de indivíduos, a definição do projeto com quais indivíduos aplicarão tratamentos em quais objetos, a escolha dos instrumentos e a avaliação da validade do experimento.
Definição da Resposta	Caso esse suporte for necessário para a realização do seu processo de experimentação, então escolha uma ferramenta que apoie a fase de Planejamento. A seguir, verifique nas Perguntas 7 e 8 quais os níveis de abstração de linguagem e de automação mais adequados ao seu processo.
<b>Pergunta 04 - No experimento, é necessário suporte de uma ferramenta que possa auxiliar na execução do experimento?</b>	

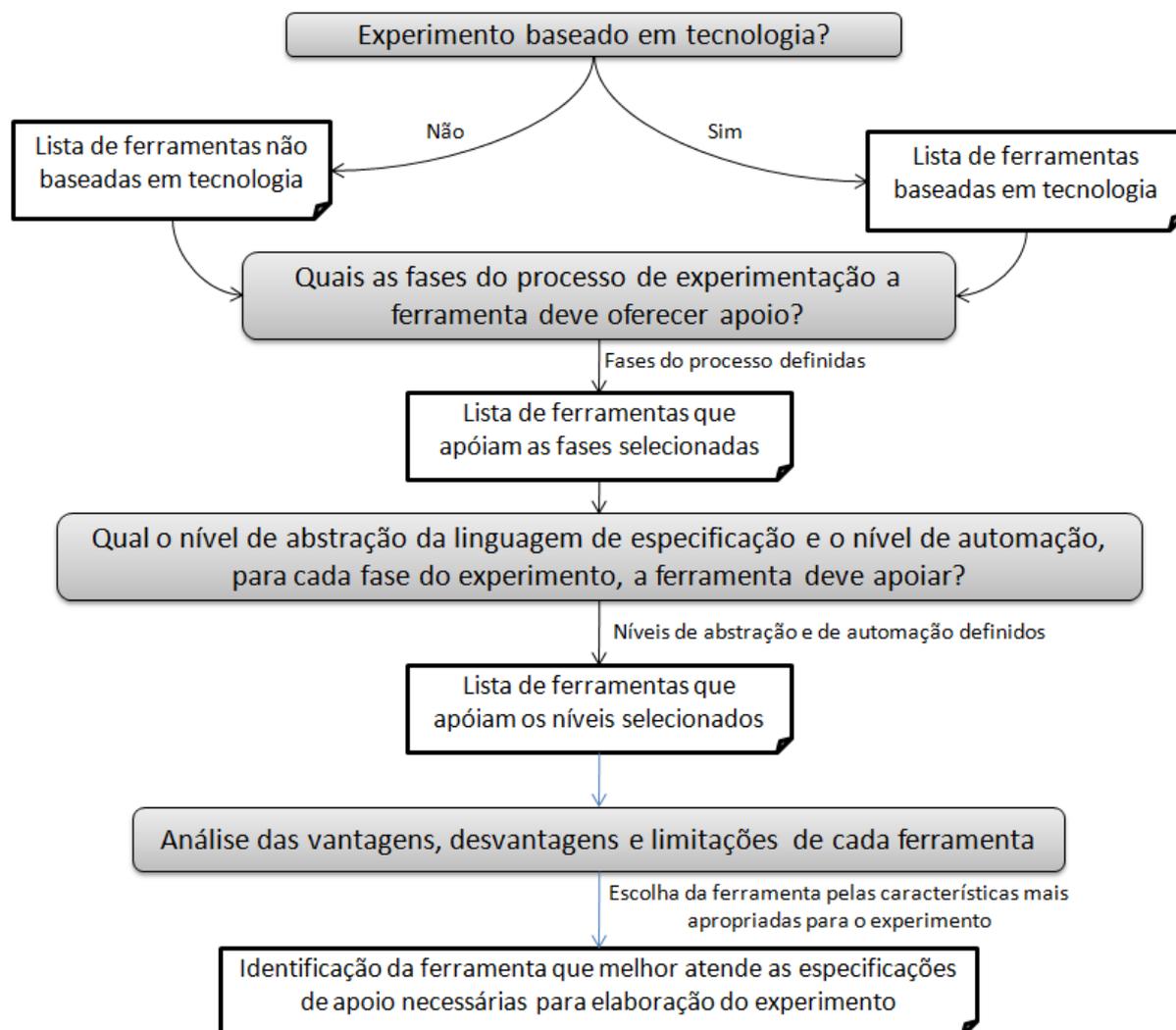
Detalhamento da pergunta	A execução é a fase em que os indivíduos aplicam os tratamentos nos objetos definidos para o experimento.
Definição da Resposta	Caso esse suporte for necessário para a realização do seu processo de experimentação, então escolha uma ferramenta que apoie a fase de Execução. A seguir, verifique nas Perguntas 7 e 8 quais os níveis de abstração de linguagem e de automação mais adequados ao seu processo.
<b>Pergunta 05 - No experimento, é necessário suporte de uma ferramenta que possa auxiliar na análise e interpretação dos resultados do experimento?</b>	
Detalhamento da pergunta	Após realizada a fase de Execução do experimento, dados foram coletados e é o momento para analisá-los e interpretá-los para gerar informações de interesse dos experimentadores.
Definição da Resposta	Caso esse suporte for necessário para a realização do seu processo de experimentação, então escolha uma ferramenta que apoie a fase de Análise e Interpretação. A seguir, verifique nas Perguntas 7 e 8 quais os níveis de abstração de linguagem e de automação mais adequados ao seu processo.
<b>Pergunta 06 - No experimento, é necessário suporte de uma ferramenta que possa auxiliar na apresentação e no empacotamento do experimento?</b>	
Detalhamento da pergunta	Assim que o experimento é concluído, há necessidade de apresentar os resultados e, se possível, empacotar todo o material produzido para futura replicação por outras pessoas. Artigos ou relatórios técnicos podem ser produzidos para a comunidade, ou os dados dos experimentos podem ser armazenados em uma base de dados corporativa, dependendo da finalidade a que foi submetido.
Definição da Resposta	Caso esse suporte for necessário para a realização do seu processo de experimentação, então escolha uma ferramenta que apoie a fase de Apresentação e Empacotamento. A seguir, verifique nas Perguntas 7 e 8 quais os níveis de abstração de linguagem e de automação mais adequados ao seu processo.

<b>Pergunta 07 - Para a fase escolhida, você prefere que a linguagem de especificação a ser utilizada para definir o experimento seja mais parecida com a linguagem natural, como a que utilizamos para escrever textos, ou mais voltada à uma linguagem de programação?</b>	
Detalhamento da pergunta	O escopo é definido através de um objeto a ser experimentado, com um determinado propósito ou intenção, buscando algum efeito com foco em qualidade, sob uma perspectiva de ponto de vista para interpretação dos resultados, e o contexto de pessoas ou tecnologias envolvidas.
Definição da Resposta	Caso deseje algo mais parecido com a linguagem natural, sua escolha deve ser por uma abstração de linguagem de especificação de Alto Nível. Caso deseje algo mais parecido com uma linguagem de programação, sua escolha deve ser por uma abstração de linguagem de especificação de Baixo Nível.
<b>Pergunta 08 - Para a fase escolhida, você prefere que a ferramenta auxilie o processo de experimentação de forma automatizada, onde ela poderá realizar as tarefas necessárias sem intervenção humana, ou a ferramenta pode auxiliar parcialmente nas tarefas mas ainda necessitando de acesso humano, ou prefere que todas as tarefas da fase sejam realizadas por humanos?</b>	
Detalhamento da pergunta	Essas tarefas podem ser, por exemplo, a geração de conteúdo, ou o gerenciamento da execução do experimento ou realizando análise de dados do experimento. Importante ressaltar que, dependendo do nível de automação escolhido, a ferramenta pode limitar mais as possibilidades de especificação nas tarefas daquela fase do experimento.
Definição da Resposta	Caso necessite de apoio totalmente automatizado, escolha uma ferramenta com suporte àquela fase de Nível Automático. Caso deseje realizar as tarefas da fase parcialmente de forma automática e manual, escolha uma ferramenta com suporte de Nível Semiautomático. E caso prefira realizar as tarefas da fase manualmente, opte por uma ferramenta com suporte a Nível Manual.

#### 4.5.2 Vantagens e Desvantagens das Ferramentas Identificadas

Sobre a análise individual de cada ferramenta, citamos a eSEE (experimental Software Engineering Environment) como a mais elaborada e com maior probabilidade de evolução. Ela compreende quatro das cinco fases do processo de experimentação. Por ser desenvolvida

Figura 4.8: Fluxograma do guia preliminar para seleção de ferramenta de apoio a um experimento



em módulos, permite a expansão de forma mais fácil. Nos estudos primários selecionados identificamos que ela suporta experimentos não baseados em tecnologia, mas se acoplado um novo módulo, a ser implementado, pode suportar também experimentos baseados em tecnologia.

A ferramenta SESE permite a execução de experimentos através de questionários web, de forma simples e rápida, mas também por essa simplicidade não provê muitos recursos para suportar diferentes tipos de experimentos. Ginger2 suporta apenas duas fases de um processo de experimentação, pois implementa apenas parcialmente o framework CAESE, sendo assim também muito limitada ao suporte oferecido a pesquisadores em Engenharia de Software. Experiment Manager Network também é direcionada a experimentos mais simples, mais especificamente àqueles executados por alunos em salas de aula, assim não recomendamos para uso em experimentos mais complexos ou de porte industrial. Por fim,

a última ferramenta sem suporte a experimentos baseados em tecnologia, a Mechanical Turk, como um software de crowdsourcing, foi utilizado no seu estudo primário para fim de realizar um experimento comparando qualidade de código fonte implementado por desenvolvedores de diferentes níveis de conhecimento, mas também se limita a este tipo de experimentação, portanto não é recomendada para a maioria dos experimentos em Engenharia de Software.

Sobre as ferramentas com apoio a experimentos baseados em tecnologia, recomendamos as duas identificadas pelo mapeamento sistemático, mas cada uma com seu propósito específico. Ginpex, sendo a melhor ferramenta para executar experimentos em ambientes computacionais com finalidade de analisar parâmetros não funcionais, como tempo de processamento, uso de memória, dentre outros. E a ferramenta Weevil, com foco em experimentos para sistemas distribuídos, utilizando ambientes de rede em larga escala.

Com a análise realizada nesta seção, identificamos os principais problemas, lacunas e vantagens entre as ferramentas de apoio a experimentos em Engenharia de Software, tanto para experimentos baseados em tecnologia ou não. Essa análise, junto com o detalhamento individual das capacidades e nível de suporte a um processo de experimentos de cada ferramenta, apresentados nas Tabelas 4.7 e 4.8, oferecem ao pesquisador interessado na elaboração de um projeto de experimentação em Engenharia de Software, um guia que os auxiliem na escolha de qual opção melhor se encaixa no experimento que desejam produzir.

Tabela 4.7: Apoio oferecido pelas ferramentas não baseadas em tecnologia

Ferramentas	Fases de um Projeto de Experimentação																			
	Escopo					Planejamento					Execução					Análise e Interpretação			Apresentação e Empacotamento	
	Nível de Abstração		Nível de Automação			Nível de Abstração		Nível de Automação			Nível de Abstração		Nível de Automação			Nível de Abstração		Nível de Automação		
	Alto	Baixo	Automático	Semiautomático	Manual	Alto	Baixo	Automático	Semiautomático	Manual	Alto	Baixo	Automático	Semiautomático	Manual	Alto	Baixo	Automático	Semiautomático	Manual
SESE	X				X	X				X	X				X					X
eSEE	X	X		X		X	X		X		X				X					X
Ginger2										X				X	X			X		
Experiment Manager										X				X	X			X		
Mechanical Turk	X				X	X	X			X				X				X		

Tabela 4.8: Apoio oferecido pelas ferramentas baseadas em tecnologia

Ferramentas	Fases de um Projeto de Experimentação																				
	Escopo					Planejamento					Execução					Análise e Interpretação			Apresentação e Empacotamento		
	Nível de Abstração		Nível de Automação			Nível de Abstração		Nível de Automação			Nível de Abstração		Nível de Automação			Nível de Abstração		Nível de Automação			
	Alto	Baixo	Automático	Semiautomático	Manual	Alto	Baixo	Automático	Semiautomático	Manual	Alto	Baixo	Automático	Semiautomático	Manual	Alto	Baixo	Automático	Semiautomático	Manual	
Ginpex	X		X		X	X	X	X		X	X		X	X		X	X	X		X	
Weevil						X			X		X					X			X		X

# Capítulo 5

## Conclusão, Trabalhos Relacionados, Ameaças a Validade e Trabalhos Futuros

Neste capítulo, apresentamos na Seção 5.1 a conclusão sobre o trabalho realizado, a Seção 5.2 detalha os trabalhos relacionados, demonstrando o valor agregado com a nossa proposta sobre os estudos existentes, a Seção 5.3 descreve as ameaças a validade desse estudo e, por fim, a Seção 5.4, com as sugestões para trabalhos futuros.

### 5.1 Conclusão

A proposta deste trabalho foi realizar um mapeamento sistemático sobre ferramentas de apoio a experimentação, e apresentar um guia para auxiliar os experimentadores em Engenharia de Software na escolha da ferramenta mais adequada ao experimento que desejam conduzir. Apresentamos no Capítulo 4 o resultado deste mapeamento sistemático e também o guia preliminar para escolha de uma ferramenta.

Utilizando como base estudos similares realizados em anos anteriores, pudemos avançar e obter novos resultados, tanto no conjunto de ferramentas ofertadas à comunidade, como observar novos aspectos sobre as ferramentas previamente identificadas. Mais cinco veículos de publicação foram utilizados na busca manual, além de obter novos resultados utilizando os mesmos veículos dos utilizados em trabalhos anteriores, mas com data atualizada até janeiro de 2018.

Com isso, além das cinco ferramentas que nosso mapeamento sistemático identificou através do trabalho de Freire et al. (2013), denominadas SESE, eSEE, Ginger2, Experiment Manager Network e Mechanical Turk, mais duas foram encontradas, a Ginpex e a Weevil. Sendo ferramentas de apoio a experimentação, a Weevil de Wang et al. (2005) tem

seu principal objetivo execução de experimentos em ambientes distribuídos, e a Ginpex de Hauck et al. (2014) que é voltada a realizar experimentos direcionados a ambientes computacionais quanto a desempenho e processamento.

Como resultado deste mapeamento sistemático, é possível verificar que as ferramentas de apoio a experimentos baseados em tecnologia são as que oferecem suporte mais automatizado em diversas fases do processo. Essas ferramentas, por tratarem de experimentos mais específicos, como análise de processamento, latência de rede, acessos distribuídos, dentre outros, tem um controle maior oferecido pela ferramenta para sua execução.

As ferramentas que não são voltadas a experimentos baseados em tecnologia, ao contrário do outro grupo, oferecem uma maior liberdade no tipo dos experimentos a que podem ser utilizadas. Para tanto oferece maior controle manual aos experimentadores, que podem customizar suas ações e adaptá-las às suas necessidades, ao custo de um esforço maior na criação do experimento.

Neste estudo, foi possível verificar que todas as ferramentas identificadas oferecem apoio a mais de uma fase de um processo de experimento, mas nenhuma suporta todas as fases definidas por Wohlin et al. (2012). A fase mais compreendida pelas ferramentas é a de Execução, sendo suportada por todas. A fase de Planejamento é a segunda com mais suporte. E a fase de Apresentação é que tem menor apoio pelas ferramentas identificadas.

Na extração de dados realizada pelo mapeamento, não identificamos ferramentas que oferecessem alguma garantia do funcionamento correto a que se propõe a atender, apresentando assim a falta confiabilidade na ferramenta. Pelos estudos primários verificados, basicamente eram apresentadas as ferramentas, suas funcionalidades e forma de execução. Uma verificação apurada da execução de experimentos utilizando essas ferramentas poderá ser efetuada em um trabalho futuro, quando comparados resultados de experimentos executados com uso dessas ferramentas e sem a utilização delas, por exemplo.

No mapeamento, foram encontradas poucas ferramentas e estudos que indicassem a aplicação das mesmas. Dentre esse conjunto identificado de ferramentas um número reduzido oferece suporte automatizado, o que indica que mais esforços podem ser realizados para elaboração de ferramentas que ofereçam automação no desenvolvimento de suas atividades, assim como compreendam todas as fases de um projeto de experimentação em Engenharia de Software.

Por fim, a análise realizada neste estudo, com o detalhamento do nível de apoio suportado em cada ferramenta, oferece a pesquisadores em Engenharia de Software um guia para ajudar na escolha de uma ferramenta de apoio a experimentos.

## 5.2 Trabalhos Relacionados

Outros dois estudos já foram realizados com temas semelhantes ao proposto aqui, porém diferem principalmente pela abrangência de veículos de publicação pesquisados, e os períodos compreendidos. Na Tabela 5.1, estão as questões de pesquisa desses estudos e a relação que cada uma tem em comparação com o trabalho proposto aqui.

O trabalho de Freire et al. (2013) foi uma revisão sistemática, com escopo de ferramentas de apoio a experimentos voltados à Engenharia de Software, e dos quinze trabalhos resultantes na pesquisa identificou sete ferramentas. Comparando com a nossa proposta de trabalho, elaboramos um mapeamento sistemático, abrangendo um escopo maior pelas questões de pesquisa definidas, principalmente pela maior cobertura sobre o apoio das ferramentas em cada fase de um projeto de experimentos. Também não avaliamos em profundidade a qualidade dos estudos, como em uma revisão sistemática. Outro diferencial do nosso trabalho é que adicionamos estudos primários de 2012 até o início de 2018, publicados após o período compreendido por Freire et al. (2013).

Em Freire et al. (2013) encontramos questões de pesquisa parecidas com as nossas. A questão QP1 da nossa proposta se assemelha às questões QP1 e QP3 de Freire et al. (2013), que identificam as ferramentas de apoio a experimentos e suas principais funcionalidades. Já as questões QP2 nos dois trabalhos têm o mesmo objetivo, pois o trabalho de Freire et al. (2013) também foi utilizado como referência para a realização deste mapeamento, porém em nossa proposta compreendemos três subquestões para cada fase apoiada pela ferramenta.

Em melhor comparação, nossas subquestões, para cada fase de um projeto de experimentos, tratam em identificar na ferramenta o grau de abstração da linguagem utilizada na especificação de experimentos, o nível de automação suportado e qual a sua confiabilidade. O grau de abstração da linguagem pode ser de alto nível, quando são utilizados modelos visuais para elaborar um projeto de experimento, por exemplo, ou de baixo nível, quando a especificação é realizada por escrita de código-fonte. O nível de automação pode ser manual, quando aquela fase do projeto necessita de intervenção total do usuário, semiautomática, quando tem uma intervenção mínima e após isso a ferramenta realiza o resto do procedimento, ou automática, quando a ferramenta executa sem necessidade de interferência humana. A confiabilidade trata de identificar as garantias oferecidas pela ferramenta quanto ao processo de experimentação, para que o usuário saiba que não ocorrerão erros no seu experimento ocasionados pela própria ferramenta utilizada.

Consideramos, para dar início a estratégia de busca por seleção manual, os mesmos estudos primários do trabalho de Freire et al. (2013), para realizar a validação do nosso protocolo de pesquisa. Pelo trabalho de Freire et al. (2013) ter um nível de granularidade maior, para que um estudo primário já selecionado pudesse ser incluir em nosso mapeamento

sistemático, era necessário que apresentasse uma ferramenta de apoio a experimentos em forma de software, e não apenas um framework para experimentação. Aplicamos os nossos critérios de seleção nestes estudos primários, e aqueles que atenderam os critérios foram incluídos em nossa lista de seleção.

Como os critérios de seleção no trabalho de referência satisfazem os critérios da nossa proposta, em uma granularidade maior, os veículos e o período que Freire et al. (2013) já realizou sua revisão sistemática não precisaram ser pesquisados novamente por nós para o mesmo período já realizado, apenas os atualizamos. Lemos os estudos primários já identificados, aplicamos os critérios de seleção e extraímos as informações que são diferentes em nossas questões de pesquisa.

Todos os cinco veículos de publicação utilizados por Freire et al. (2013) na busca manual, entre janeiro de 2002 até dezembro de 2011, foram atualizados até a janeiro de 2018. Também incluímos outros cinco veículos compreendendo o período desde janeiro de 2002 até janeiro de 2018, que aumentaram significativamente o número de publicações pesquisadas. A Tabela 4.1 contém a relação com todos os veículos pesquisados e os períodos compreendidos.

Tabela 5.1: Questões de pesquisa X trabalhos relacionados

Esta Dissertação	Freire et al. (2013)	
QP1	QP1	Quais ferramentas foram propostas para suportar experimentos controlados em Engenharia de Software?
QP1	QP3	Quais as principais funcionalidades suportadas pelas ferramentas propostas?
QP2	QP2	Quais fases de experimentos controlados as ferramentas propostas dão suporte?
	Ferreira et al. (2017)	
QP2	QP1	Como a abordagem de Orientação a Modelos, para apoiar experimentos individuais em Engenharia de Software, seguem os guias de Projetos de experimentos?

O trabalho de Ferreira et al. (2017) teve o escopo direcionado à metodologia de Orientação a Modelos, quando ela é utilizada para auxiliar projetos de experimentos em Engenharia de Software. O trabalho estabeleceu oito critérios detalhados, mas que tratam apenas as fases de Escopo e de Planejamento. Comparando com a proposta do nosso estudo, buscamos verificar todas as fases de um projeto de experimentos, sendo assim

mais abrangente que o trabalho de Ferreira et al. (2017), além de que não restringimos os estudos primários especificamente a trabalhos da metodologia de Orientação a Modelos.

Em sua questão de pesquisa, o trabalho de Ferreira et al. (2017) tem mais semelhança com a QP2 do nosso mapeamento sistemático por tratar mais especificamente de fases dos projetos de experimentos, conforme observado na Tabela 5.1.

Dessa forma, o estudo proposto por nossa pesquisa traz uma nova abordagem para um mapeamento de ferramentas de experimentação em Engenharia de Software, compreendendo mais detalhes sobre as fases de um processo de experimentos, e contemplando outros fatores sobre a ferramenta como produto, diferenciando-se de outros trabalhos já publicados, relacionados neste capítulo.

### 5.3 Ameaças à Validade do Estudo

As ameaças a validade desse estudo incluem, dentre elas, que mais veículos de publicações sobre engenharia de software poderiam ser pesquisados, na expectativa de encontrar mais ferramentas de apoio a experimentos. Porém, devido à limitação natural de tempo e recursos humanos, não foi possível pesquisar em um número maior de veículos. Como exemplo, dois veículos que não realizamos a pesquisa foi o European Association of Science Editors - EASE, evento que tem várias edições com assuntos relacionados ao nosso trabalho, e o Information and Software Technology - IST, que publica muitos estudos secundários relevantes.

Outro fator de validade é pela pesquisa manual que foi realizada para identificar estudos primários que apresentassem ferramentas, pois um erro humano pode ter ocorrido na leitura do resumo de um estudo, e assim deixado de incluí-lo na lista de leitura completa. Da mesma forma, a extração dos dados de um estudo selecionado foi realizada de forma manual, e que pode apresentar algum equívoco na interpretação das informações constantes nos estudos. Os autores das ferramentas não foram contatados, o que poderia deixar a pesquisa mas precisa quanto às funcionalidades e capacidades das ferramentas. Mais uma ameaça a validade apresentada em nosso estudo é o guia preliminar na última seção do Capítulo 4, que apenas foi descrito mas não verificado sua usabilidade em algum projeto real.

O último item que descrevemos como uma ameaça a validade deste trabalho de mapeamento sistemático, se refere ao critério de inclusão, definido na metodologia de pesquisa do Capítulo 3. Por apenas selecionarmos estudos primários que definissem a utilização de uma ferramenta de apoio a experimentos no título ou no resumo do estudo, pode haver trabalhos que não foram incluídos para leitura completa, mas que talvez descrevessem novas ferramentas no interior do seu conteúdo completo. Entretanto, centenas,

senão milhares de estudos apenas citam realizar experimentos, e para realizar a leitura integral desses estudos em busca de uma nova proposta de ferramenta, exigiria um enorme esforço e maior quantidade de recursos humanos disponíveis.

## 5.4 Trabalhos Futuros

Sobre trabalhos futuros, podemos sugerir abordar os itens que foram identificados em nossas ameaças à validade. Realizar mapeamentos e revisões sistemáticas por um número maior de integrantes reduz o risco de um estudo primário, que apresente uma ferramenta de apoio a experimentos, deixe de ser selecionado, ou que tenha seus dados extraídos de forma equivocada.

Considerar mais veículos de publicação e sem restrição de data de início da busca, pode trazer novas ferramentas que não foram identificadas neste trabalho. Realizar a técnica de snowballing nos estudos primários deste mapeamento e dos trabalhos de Freire et al. (2013) e de Ferreira et al. (2017), que consiste em realizar uma busca manual inversa utilizando como base as referências desses estudos, pode trazer mais resultados do que apenas fazendo a busca diretamente nos veículos de publicação.

Sugerimos também a realização de um estudo de usabilidade sobre o guia preliminar, apresentado no Capítulo 4, para validar sua eficácia em auxiliar um pesquisador em Engenharia de Software na escolha da ferramenta mais adequada para seu experimento.

Também há possibilidade de realizar mais estudos onde a aplicação das ferramentas seja em projetos reais de experimentos, com isso aumentará o rigor e a relevância industrial sobre essas ferramentas, e pode estimular ao mercado na sua utilização. Um estudo de usabilidade das ferramentas identificadas em novos experimentos, e em replicações de experimentos originais para validar os resultados, pode apresentar mais pontos fortes e falhas que não foram identificados neste mapeamento.

Evoluir as ferramentas existentes, encontradas neste mapeamento sistemático, também é uma oportunidade. Pontos como a cobertura da fase de Apresentação, ou a demonstração de maior confiabilidade na ferramenta, são assuntos para novas pesquisas e publicações de novos materiais sobre essas ferramentas. Destacamos que, em nosso grupo de pesquisa, estamos projetando uma ferramenta que suporta automação de experimentos baseados em tecnologia cobrindo todas as fases do processo de experimentação.

# Referências

- Arisholm, E., D. I. K. Sjøberg, G. Carelius, e Y. Lindsjörn  
2002a. Sese - an experiment support environment for evaluating software engineering technologies. *Nordic Workshop on Programming and Software Development Tools and Techniques*. 24
- Arisholm, E., D. I. K. Sjøberg, G. Carelius, e Y. Lindsjörn  
2002b. A web-based support environment for software engineering experiments. *Nordic Journal of Computing*, 9:231–247. 24
- Chapetta, W. A., P. S. M. Santos, e G. H. Travassos  
2005. Supporting meta-description activities in experimental software engineering environments. *Experimental Software Engineering Latin American Workshop*. 25
- Cook, T. e D. Campbell  
1979. *Quasi-experimentation: design & analysis issues for field settings*. Rand McNally College. 9
- Easterbrook, S., J. Singer, M. Storey, e D. Damian  
2008. *Selecting Empirical Methods for Software Engineering Research*, Pp. 285–311. London: Springer London. 35
- Ferreira, W., M. T. Baldassarre, S. Soares, B. Cartaxo, e G. Visaggio  
2017. A comparative study of model-driven approaches for scoping and planning experiments. *Proceedings of the 21st International Conference on Evaluation and Assessment in Software Engineering*, Pp. 78–87. 54, 55, 56
- Freire, M. A., D. A. da Costa, E. C. Neto, T. Medeiros, U. Kulesza, E. Aranha, e S. Soares  
2013. Automated support for controlled experiments in software engineering: A systematic review (S). *The 25th International Conference on Software Engineering and Knowledge Engineering, Boston, MA, USA, June 27-29, 2013.*, Pp. 504–509. 1, 2, 13, 17, 22, 23, 24, 26, 51, 53, 54, 56, 62, 63, 65, 66, 68, 70, 72
- Hauck, M., M. Kuperberg, N. Huber, e R. Reussner  
2014. Deriving performance-relevant infrastructure properties through model-based experiments with ginpeX. *Software & Systems Modeling*, 13(4):1345–1365. 25, 52
- Hochstein, L., T. Nakamura, F. Shull, N. Zazworka, V. R. Basili, e M. Zelkowitz  
2008. An environment for conducting families of software engineering experiments. *Advances in Computers*, 74:175–200. 25

- Jedlitschka, A. e D. Pfahl  
2005. Reporting guidelines for controlled experiments in software engineering. In *2005 International Symposium on Empirical Software Engineering, 2005.*, Pp. 10 pp.–. 11
- Juristo, N. e A. M. Moreno  
2013. *Basics of software engineering experimentation*. Springer Science & Business Media. 1, 6, 35
- Juristo, N. J. e A. M. Moreno  
2001. *Basics of software engineering experimentation*. Kluwer Academic Press. 5
- Kitchenham, B. e S. Charters  
2007. Guidelines for performing systematic literature reviews in software engineering. 2. 1, 3, 4, 11, 12, 13
- Kitchenham, B. A., P. Brereton, M. Turner, M. K. Niazi, S. Linkman, R. Pretorius, e D. Budgen  
2010. Refining the systematic literature review process—two participant-observer case studies. *Empirical Software Engineering*, 15(6):618–653. 17
- Mian, P. G., G. H. Travassos, e A. R. C. Rocha  
2004a. eese: a computerized infrastructure for experimental software engineering. *Experimental Software Engineering Latin American Workshop*. 25
- Mian, P. G., G. H. Travassos, e A. R. C. Rocha  
2005. A computerized infrastructure for supporting experimentation in software engineering. *Experimental Software Engineering Latin American Workshop*. 25
- Mian, P. G., G. H. Travassos, A. R. C. Rocha, e A. C. C. Natali  
2004b. Towards a computerized infrastructure for managing experimental software engineering knowledge. *Jornadas Iberoamericanas em Engenharia del Software e Ingeniería del Conocimiento*. 25
- Neto, A. C. D., R. F. Barcelos, W. A. Chapetta, P. S. M. Santos, e S. N. Mafra  
2004. Infrastructure for se experiments definition and planning. *Experimental Software Engineering Latin American Workshop*. 24
- Petersen, K., R. Feldt, S. Mujtaba, e M. Mattsson  
2008. Systematic mapping studies in software engineering. *Proceedings of the 12th international conference on Evaluation and Assessment in Software Engineering*, 8:68–77. 11, 12
- Sjoberg, D. I., A. Bente, E. Arisholm, T. Dyba, M. Jorgensen, A. Karahasanovic, E. F. Koren, e M. Vokac  
2002. Conducting realistic experiments in software engineering. In *Proceedings International Symposium on Empirical Software Engineering*, Pp. 17–26. 24
- Stolee, K. T. e S. Elbaum  
2010. Exploring the use of crowdsourcing to support empirical studies in software engineering. *Proceedings of the 2010 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement*, Pp. 35:1–35:4. 25, 29

- Torii, K., K. Matsumoto, K. Nakakoji, Y. Takada, S. Takada, e K. Shima  
1999. Ginger2: an environment for computer-aided empirical software engineering. *IEEE Transactions on Software Engineering*, 25(4):474–492. 25
- Travassos, G. H., S. S. Medeiros, P. G. Mian, A. C. D. Neto, e J. C. d. A. Biolchini  
2008. An environment to support large scale experimentation in software engineering. *13th IEEE International Conference on Engineering of Complex Computer Systems (iceccs 2008)*, Pp. 193–202. 25, 36
- Vale, T., E. Santana de Almeida, V. Alves, U. Kulesza, N. Niu, e R. de Lima  
2017. Software product lines traceability: A systematic mapping study. *Information and Software Technology*, 84(Supplement C):1 – 18. 12, 13
- Wang, Y., M. J. Rutherford, A. Carzaniga, e A. L. Wolf  
2005. Automating experimentation on distributed testbeds. *Proceedings of the 20th IEEE/ACM International Conference on Automated Software Engineering*, Pp. 164–173. 25, 51
- Wohlin, C., P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, e A. Wesslén  
2012. *Experimentation in Software Engineering*. Springer Science & Business Media. 1, 2, 3, 6, 7, 8, 30, 35, 52

# Anexo I

## Formulário de Extração de Dados

### Dados Básicos

Dados do Estudo Primário	
Título	
Autores	
Local de Publicação	
Ano de Publicação	
Comentários (opcional)	

### Fases de Experimentação Cobertas pela Ferramenta

Fase do Experimento	Oferece Apoio	Nível da Linguagem	Nível da Automação	Oferece Garantia
Escopo				
Planejamento				
Execução				
Análise e Interpretação				
Apresentação				

### Técnicas Utilizadas pela Ferramenta

1. Nome da abordagem utilizada para realizar a experimentação (se existente):

2. Forma de execução (funcionamento) e principais características:
3. O tipo de experimento apoiado pela ferramenta é baseado em tecnologia ou não?

# Anexo II

## Formulários Preenchidos pela Extração de Dados

### Formulário de Extração de Dados 01

#### Dados Básicos

<b>Dados do Estudo Primário</b>	
Título	Conducting realistic experiments in software engineering
Autores	Sjoberg, D. I., A. Bente, E. Arisholm, T. Dyba, M. Jorgensen, A. Karahasanovic, E. F. Koren, e M. Vokac
Local de Publicação	Proceedings International Symposium on Empirical Software Engineering
Ano de Publicação	2002
Comentários (opcional)	Páginas 17 à 26. Este estudo primário foi identificado no trabalho de Freire et al. (2013) e é um dos três estudos que fazem referência à ferramenta SESE

<b>Dados do Estudo Primário</b>	
Título	SESE – an Experiment Support Environment for Evaluating Software Engineering Technologies
Autores	Arisholm, E., D. I. K. Sjøberg, G. Carelius, e Y. Lindsjörn
Local de Publicação	Workshop on Programming and Software Development Tools and Techniques.
Ano de Publicação	2002

Comentários (opcional)	Este estudo primário foi identificado no trabalho de Freire et al. (2013) e é um dos três estudos que fazem referência à ferramenta SESE
------------------------	--

<b>Dados do Estudo Primário</b>	
Título	A Web-based Support Environment for Software Engineering Experiments
Autores	Arisholm, E., D. I. K. Sjøberg, G. Carelius, e Y. Lindsjörn
Local de Publicação	Nordic Journal of Computing
Ano de Publicação	2002
Comentários (opcional)	Páginas 231 à 247. Este estudo primário foi identificado no trabalho de Freire et al. (2013) e é um dos três estudos que fazem referência à ferramenta SESE

## Fases de Experimentação Cobertas pela Ferramenta

Fase do Experimento	Oferece Apoio	Nível da Linguagem	Nível da Automação	Oferece Garantia
<b>Escopo</b>	Sim	Alto	Manual	Não
<b>Planejamento</b>	Sim	Alto	Manual	Não
<b>Execução</b>	Sim	Alto	Manual	Não
<b>Análise e Interpretação</b>	Não	-	-	-
<b>Apresentação</b>	Sim	Alto	Semiautomático	Não

## Técnicas Utilizadas pela Ferramenta

1. Nome da abordagem utilizada para realizar a experimentação (se existente)

R: Web-based Simula Experiment Support Environment (SESE)

2. Forma de execução (funcionamento) e principais características

R: A ferramenta SESE realiza a criação de formulários Web com questões referentes ao experimento desejado. Experimentadores em Engenharia de Software definem os questionários detalhados, descrevendo suas tarefas e incluindo os códigos necessários. Participantes preenchem a sequência de questões e submetem

suas respostas, que são coletadas pela ferramenta. Cada sessão da execução de um experimento pode ser monitorada pelo experimentador, o que permite suporte a experimentos de larga escala. Por fim, o sistema pode gerar gráficos de acordo com o resultado do experimento.

3. O tipo de experimento apoiado pela ferramenta é baseado em tecnologia ou não?

R: A ferramenta SESE não suporta experimentos baseados em tecnologia

## Formulário de Extração de Dados 02

### Dados Básicos

<b>Dados do Estudo Primário</b>	
Título	Towards a Computerized Infrastructure for Managing Experimental Software Engineering Knowledge
Autores	Mian, P. G., Travassos, G. H., Rocha, A. R. C., e Natali, A. C. C.
Local de Publicação	Jornadas Iberoamericanas em Ingeniería del Software e Ingeniería del Conocimiento
Ano de Publicação	2004
Comentários (opcional)	Este estudo primário foi identificado no trabalho de Freire et al. (2013) e é um dos seis estudos que fazem referência à ferramenta eSEE

<b>Dados do Estudo Primário</b>	
Título	Infrastructure for SE Experiments Definition and Planning
Autores	Neto, A. C. D., R. F. Barcelos, W. A. Chapetta, P. S. M. Santos, e S. N. Mafra
Local de Publicação	Experimental Software Engineering Latin American Workshop
Ano de Publicação	2004
Comentários (opcional)	Este estudo primário foi identificado no trabalho de Freire et al. (2013) e é um dos seis estudos que fazem referência à ferramenta eSEE

<b>Dados do Estudo Primário</b>	
Título	eSEE: a Computerized Infrastructure for Experimental Software Engineering
Autores	Mian, P. G., Travassos, G. H., Rocha, A. R. C., e Natali, A. C. C.
Local de Publicação	Experimental Software Engineering Latin American Workshop
Ano de Publicação	2004
Comentários (opcional)	Este estudo primário foi identificado no trabalho de Freire et al. (2013) e é um dos seis estudos que fazem referência à ferramenta eSEE

<b>Dados do Estudo Primário</b>	
Título	A computerized infrastructure for supporting experimentation in software engineering
Autores	Mian, P. G., G. H. Travassos, e A. R. C. Rocha
Local de Publicação	Experimental Software Engineering Latin American Workshop
Ano de Publicação	2005
Comentários (opcional)	Este estudo primário foi identificado no trabalho de Freire et al. (2013) e é um dos seis estudos que fazem referência à ferramenta eSEE

<b>Dados do Estudo Primário</b>	
Título	Supporting Meta-Description Activities in Experimental Software Engineering Environments
Autores	Chapetta, W. A., P. S. M. Santos, e G. H. Travassos
Local de Publicação	Experimental Software Engineering Latin American Workshop
Ano de Publicação	2005
Comentários (opcional)	Este estudo primário foi identificado no trabalho de Freire et al. (2013) e é um dos seis estudos que fazem referência à ferramenta eSEE

<b>Dados do Estudo Primário</b>	
Título	An environment to support large scale experimentation in software engineering
Autores	Travassos, G. H., S. S. Medeiros, P. G. Mian, A. C. D. Neto, e J. C. d. A. Biolchini
Local de Publicação	International Conference on Engineering of Complex Computer Systems
Ano de Publicação	2008
Comentários (opcional)	Este estudo primário foi identificado no trabalho de Freire et al. (2013) e é um dos seis estudos que fazem referência à ferramenta eSEE

## Fases de Experimentação Cobertas pela Ferramenta

Fase do Experimento	Oferece Apoio	Nível da Linguagem	Nível da Automação	Oferece Garantia
Escopo	Sim	Alto/Baixo	Semiautomático	Não
Planejamento	Sim	Alto/Baixo	Semiautomático	Não
Execução	Sim	Alto	Manual	Não
Análise e Interpretação	Não	-	-	-
Apresentação	Sim	Alto	Semiautomático	Não

## Técnicas Utilizadas pela Ferramenta

1. Nome da abordagem utilizada para realizar a experimentação (se existente)

R: experimental Software Engineering Environment (eSEE)

2. Forma de execução (funcionamento) e principais características

R: Esta ferramenta permite criar a infraestrutura para um ambiente experimental em Engenharia de Software, seguindo um modelo para empacotamento de experimentos, que define uma taxonomia de documento para representar a informação necessária a execução de experimentos. o eSEE é dividido em três níveis de abstração. O primeiro nível de abstração é o Nível Meta, que contém conhecimento comum sobre engenharia de software experimental. O segundo nível, de Configuração, trata sobre os tipos de estudos experimentais e sua forma de conhecimento. O terceiro nível, de Execução, é o conhecimento sobre um determinado estudo. Essa infraestrutura é definida para suportar processos de experimentação em larga escala, fornecendo recursos para que experimentadores em Engenharia de Software, em locais diferentes, possam realizar experimentos através de uma rede de computadores.

3. O tipo de experimento apoiado pela ferramenta é baseado em tecnologia ou não?

R: A ferramenta eSEE não suporta experimentos baseados em tecnologia

# Formulário de Extração de Dados 03

## Dados Básicos

Dados do Estudo Primário	
Título	Exploring the Use of Crowdsourcing to Support Empirical Studies in Software Engineering
Autores	Stolee, K. T. and Elbaum, S.
Local de Publicação	ACM-IEEE International Symposium on Empirical Software Engineering and Measurement
Ano de Publicação	2010
Comentários (opcional)	Este estudo primário foi identificado no trabalho de Freire et al. (2013)

## Fases de Experimentação Cobertas pela Ferramenta

Fase do Experimento	Oferece Apoio	Nível da Linguagem	Nível da Auto- mação	Oferece Garantia
Escopo	Sim	Alto	Manual	Não
Planejamento	Sim	Alto/Baixo	Manual	Não
Execução	Sim	Alto	Manual	Não
Análise e Interpretação	Não	-	-	-
Apresentação	Não	-	-	-

## Técnicas Utilizadas pela Ferramenta

1. Nome da abordagem utilizada para realizar a experimentação (se existente)

R: Mechanical Turk

2. Forma de execução (funcionamento) e principais características

R: Mechanical Turk é um software de crowdsourcing, que pode ser utilizado para obter resultados sobre um experimento envolvendo muitos indivíduos, sendo eles

os usuários desse software. A ferramenta permite selecionar o tipo e o número de indivíduos a executar o experimento, e também gerenciar esses usuários conforme o experimento estiver sendo conduzido.

3. O tipo de experimento apoiado pela ferramenta é baseado em tecnologia ou não?

R: A ferramenta Mechanical Turk não suporta experimentos baseados em tecnologia

# Formulário de Extração de Dados 04

## Dados Básicos

Dados do Estudo Primário	
Título	An Environment for Conducting Families of Software Engineering Experiments
Autores	Hochstein, L., Nakamura, T., Shull, F., Zazworka, N., R. Basili, V., e Zelkowitz, M.
Local de Publicação	Advances in Computers
Ano de Publicação	2008
Comentários (opcional)	Páginas 175 à 200. Este estudo primário foi identificado no trabalho de Freire et al. (2013)

## Fases de Experimentação Cobertas pela Ferramenta

Fase do Experimento	Oferece Apoio	Nível da Linguagem	Nível da Automação	Oferece Garantia
Escopo	Não	-	-	-
Planejamento	Não	-	-	-
Execução	Sim	Alto	Manual	Não
Análise e Interpretação	Sim	Alto	Semiautomático	Não
Apresentação	Não	-	-	-

## Técnicas Utilizadas pela Ferramenta

1. Nome da abordagem utilizada para realizar a experimentação (se existente)

R: Experiment Manager framework

2. Forma de execução (funcionamento) e principais características

R: Experiment Manager é uma infraestrutura, formada por um conjunto de ferramentas, que simplifica o processo de coletar, gerenciar e sanitizar dados de experimentos

em engenharia de software, executados em salas de aula, e minimiza a mudança comportamental natural dos indivíduos relacionados ao experimento. O ambiente é um conjunto de ferramentas com suporte a experimentos em Engenharia de Software, que envolvam tarefas de desenvolvimento e fluxo de trabalho. Essas ferramentas também realizam a coleta e análise de dados dos experimentos aplicados em salas de aula.

3. O tipo de experimento apoiado pela ferramenta é baseado em tecnologia ou não?

R: A ferramenta Experiment Manager framework não suporta experimentos baseados em tecnologia

# Formulário de Extração de Dados 05

## Dados Básicos

Dados do Estudo Primário	
Título	Ginger2: an environment for computer-aided empirical software engineering
Autores	Torii, K., Matsumoto, K., Nakakoji, K., Takada, Y., Takada, S., e Shima, K.
Local de Publicação	IEEE Transactions on Software Engineering
Ano de Publicação	1999
Comentários (opcional)	Páginas 474 à 492. Este estudo primário foi identificado no trabalho de Freire et al. (2013)

## Fases de Experimentação Cobertas pela Ferramenta

Fase do Experimento	Oferece Apoio	Nível da Linguagem	Nível da Automação	Oferece Garantia
Escopo	Não	-	-	-
Planejamento	Não	-	-	-
Execução	Sim	Alto	Manual	Não
Análise e Interpretação	Sim	Alto	Semiautomático	Não
Apresentação	Não	-	-	-

## Técnicas Utilizadas pela Ferramenta

1. Nome da abordagem utilizada para realizar a experimentação (se existente)

R: Ginger2

2. Forma de execução (funcionamento) e principais características

R: O framework CAESE, utilizado pelo Ginger2, cobre todas as fases de um projeto de experimentação, porém a ferramenta implementa apenas parte do framework,

oferecendo apoio para as fases de Execução e Análise e Interpretação.

3. O tipo de experimento apoiado pela ferramenta é baseado em tecnologia ou não?

R: A ferramenta Ginger2 não suporta experimentos baseados em tecnologia

# Formulário de Extração de Dados 06

## Dados Básicos

Dados do Estudo Primário	
Título	Deriving performance-relevant infrastructure properties through model-based experiments with Ginpex
Autores	Hauck, M., Kuperberg, M., Huber, N., and Reussner, R.
Local de Publicação	Software & Systems Modeling
Ano de Publicação	2014
Comentários (opcional)	Páginas 1345 à 1365

## Fases de Experimentação Cobertas pela Ferramenta

Fase do Experimento	Oferece Apoio	Nível da Linguagem	Nível da Automação	Oferece Garantia
Escopo	Sim	Alto	Manual/Autom.	Não
Planejamento	Sim	Alto/Baixo	Manual/Autom.	Não
Execução	Sim	Alto	Semiaut./Autom.	Não
Análise e Interpretação	Sim	Alto/Baixo	Manual/Autom.	Não
Apresentação	Não	-	-	-

## Técnicas Utilizadas pela Ferramenta

1. Nome da abordagem utilizada para realizar a experimentação (se existente)

R: Goal-oriented INfrastructure Performance EXperiments (Ginpex)

2. Forma de execução (funcionamento) e principais características

R: Ginpex é uma ferramenta direcionada a experimentos baseados em tecnologia, que verificam parâmetros como nível de performance de um sistema, tempo de processamento, uso de recursos computacionais, dentre outros. Utilizando uma abordagem orientada a objetivos, que por meio de modelos, gera experimentos exe-

cutáveis para detectar e quantificar automaticamente propriedades de infraestrutura que afetam o desempenho. Através de execução automatizada e suporte à análise e dados, auxilia os experimentadores a analisarem propriedades de infraestrutura relevantes para o desempenho. A ferramenta oferece Um conjunto de modelos de experimentos predefinidos, que além de sua execução também realiza a avaliação dos resultados.

3. O tipo de experimento apoiado pela ferramenta é baseado em tecnologia ou não?

R: A ferramenta Ginpex suporta experimentos baseados em tecnologia

# Formulário de Extração de Dados 07

## Dados Básicos

Dados do Estudo Primário	
Título	Automating Experimentation on Distributed Testbeds
Autores	Wang, Y., Rutherford, M. J., Carzaniga, A., and Wolf, A. L.
Local de Publicação	IEEE/ACM International Conference on Automated Software Engineering
Ano de Publicação	2005
Comentários (opcional)	Páginas 164 à 173

## Fases de Experimentação Cobertas pela Ferramenta

Fase do Experimento	Oferece Apoio	Nível da Linguagem	Nível da Automação	Oferece Garantia
Escopo	Não	-	-	-
Planejamento	Sim	Alto	Semiautomático	Não
Execução	Sim	Alto	Automático	Não
Análise e Interpretação	Sim	Alto	Semiautomático	Não
Apresentação	Sim	Alto	Semiautomático	Não

## Técnicas Utilizadas pela Ferramenta

1. Nome da abordagem utilizada para realizar a experimentação (se existente)

R: Weevil

2. Forma de execução (funcionamento) e principais características

R: A ferramenta Weevil tem como principal característica a execução de experimentos em sistemas distribuídos, usando preferencialmente redes computacionais chamadas Testbeds, próprias para avaliação de softwares que necessitam controlar acessos remotos e troca de dados entre diversas máquinas clientes. Utiliza uma abor-

dagem baseada em modelos, definindo em um conjunto desses modelos o sistema em experimentação, qual a rede computacional que será utilizada, e os comportamentos dos clientes que conduzem o experimento. Com base nessa configuração, oferece execução automatizada.

3. O tipo de experimento apoiado pela ferramenta é baseado em tecnologia ou não?

R: A ferramenta Weevil suporta experimentos baseados em tecnologia