

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL  
FACULDADE DE INFORMÁTICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**SISTEMATIZAÇÃO DAS EVIDÊNCIAS  
EMPÍRICAS EM DESENVOLVIMENTO  
DISTRIBUÍDO DE SOFTWARE**

**ANTÔNIO RAFAEL DA ROSA TECHIO**

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Ciência da Computação na Pontifícia Universidade Católica do Rio Grande do Sul.

Orientador: Prof. Dr. Rafael Prikladnicki

Porto Alegre  
2014



T255s      Techio, Antônio Rafael da Rosa  
              Sistematização das evidências empíricas em desenvolvimento  
              distribuído de software / Antônio Rafael da Rosa Techio. – Porto  
              Alegre, 2014.  
              151 p.  
  
              Diss. (Mestrado) – Fac. de Informática, PUCRS.  
              Orientador: Prof. Dr. Rafael Prikladnicki.  
  
              1. Informática. 2. Engenharia de Software. 3. Sistemas  
              Distribuídos. 4. Taxionomia. 5. I. Prikladnicki, Rafael. II. Título.

CDD 005.1

**Ficha Catalográfica elaborada pelo  
Setor de Tratamento da Informação da BC-PUCRS**





Pontifícia Universidade Católica do Rio Grande do Sul  
FACULDADE DE INFORMÁTICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

## TERMO DE APRESENTAÇÃO DE DISSERTAÇÃO DE MESTRADO

Dissertação intitulada "Sistematização das Evidências Empíricas em Desenvolvimento Distribuído de Software" apresentada por Antônio Rafael da Rosa Techio como parte dos requisitos para obtenção do grau de Mestre em Ciência da Computação, aprovada em 17/03/2014 pela Comissão Examinadora:

Prof. Dr. Rafael Prikladnicki –  
Orientador

PPGCC/PUCRS

Prof. Dr. Paulo Henrique Lemelle Fernandes –

PPGCC/PUCRS

Profa. Dra. Tayana Uchôa Conte –

UFAM

Homologada em...../...../....., conforme Ata No. .... pela Comissão Coordenadora.

Prof. Dr. Luiz Gustavo Leão Fernandes  
Coordenador.

**PUCRS**

**Campus Central**  
Av. Ipiranga, 6681 – P32 – sala 507 – CEP: 90619-900  
Fone: (51) 3320-3611 – Fax (51) 3320-3621  
E-mail: [ppgcc@pucrs.br](mailto:ppgcc@pucrs.br)  
[www.pucrs.br/facin/pos](http://www.pucrs.br/facin/pos)



## DEDICATÓRIA

*“Para meus pais, Antônio e Lizete, e para minha esposa, Vanessa,  
por sempre acreditarem e me incentivarem no alcance dos meus objetivos”.*



## AGRADECIMENTOS

Primeiramente gostaria de agradecer a Deus, por estar sempre ao meu lado, abençoando cada etapa da minha vida.

Gostaria de agradecer especialmente a minha esposa, Vanessa, pelo apoio e companheirismo, sem você eu não chegaria até aqui. Obrigado pelo carinho, amor e principalmente pela compreensão durante esses dois anos que não foram nada fáceis.

À minha família, por me apoiarem em todas as minhas escolhas e por estarem sempre ao meu lado nos momentos mais difíceis e também nos momentos de alegria.

Ao meu orientador, Prof. Dr. Rafael Prikladnicki, pela oportunidade, confiança e pelas críticas e sugestões que tanto ajudaram no decorrer da pesquisa. Muito obrigado.

Ao Prof. Dr. Paulo Fernandes, por suas críticas, comentários e sugestões dadas em todas as avaliações intermediárias desta pesquisa, visando sempre contribuir.

Aos meus colegas de mestrado pela ajuda e amizade durante essa caminhada de dois anos de muito trabalho.

A PUCRS, Faculdade de Informática, membros do PPGCC e professores, pela excelente infraestrutura disponível e por proporcionar um aprendizado de qualidade na área acadêmica.

A Dell *Inc.*, empresa onde eu trabalho, pelo apoio e flexibilidade necessária para conduzir esse trabalho.

A todos entrevistados pelo tempo dispensado e pela rica contribuição.

Enfim, a todos aqueles que contribuíram de alguma forma para a conclusão deste mestrado, meu muito obrigado.



# SISTEMATIZAÇÃO DAS EVIDÊNCIAS EMPÍRICAS EM DESENVOLVIMENTO DISTRIBUÍDO DE SOFTWARE

## RESUMO

O Desenvolvimento Distribuído de Software (DDS) é recente como área de pesquisa e vem crescendo ao longo dos últimos anos. Como resultado, um grande número de pesquisas e estudos na indústria foi publicado. No entanto, devido a grande quantidade de termos envolvidos nesse cenário, muitas vezes um conceito é utilizado de forma equivocada, tornando o processo de busca por evidências empíricas uma tarefa complexa. Os cenários de DDS são diversos e o que se aplica em um contexto talvez não se aplique diretamente em outro. Dessa forma é necessário entender como os resultados empíricos em DDS podem ser relatados para ser útil para profissionais e pesquisadores desta área. A pesquisa empírica na área de DDS poderia ser refinada se houvesse uma maneira sistemática para identificar as diferentes dimensões e características para o qual se aplica a evidência empírica. Por esta razão, o objetivo desta dissertação de mestrado é identificar que tipo de informação de contexto é relevante, para pesquisadores e profissionais da indústria, em estudos com evidências empíricas e propor uma forma de sistematização das evidências empíricas em DDS, visando contribuir, entre outras coisas, com o desenvolvimento de novas pesquisas nesta área. Para o desenvolvimento desta pesquisa utilizou-se métodos secundários (revisão sistemática da literatura, do tipo mapeamento sistemático) e primários (*survey* com especialistas) de pesquisa. Este estudo contribui no sentido de propor uma taxonomia para sistematizar as evidências empíricas geradas a partir das pesquisas científicas conduzidas na área de DDS.

**Palavras Chave:** Sistematização do conhecimento, Desenvolvimento Distribuído de Software, Taxonomia, Evidência Empírica, Estudos Empíricos.



# SYSTEMATIZATION OF EMPIRICAL EVIDENCE IN DISTRIBUTED SOFTWARE DEVELOPMENT

## ABSTRACT

Distributed Software Development (DSD) is a recent area and has been growing over the last few years. As a result, a quite number of research and industrial studies have been published. However, due to the large number of terms involved in this scenario, a concept is often used wrongly, making the process of searching for empirical evidence a complex task. The DSD scenarios are diverse and what works in one context might not directly apply in another. Thus it is necessary to understand how DSD related empirical findings should be reported to be useful for practitioners and researchers. Empirical research in DSD could be refined if there was a systematic way to identify the different dimensions and characteristics for which empirical evidence applies. For this reason, the goal of this dissertation is to identify the type of context information that is relevant to researchers and practitioners in studies with empirical evidence and propose a way to systematize the empirical evidence in DSD, in order to contribute, among other things, to the future development of research in the area. For the development of this research we have used both secondary (systematic mapping study) and primary (expert opinion survey) research methods. The main contribution of this research is the taxonomy to systematize the empirical evidences generated from scientific research conducted in the DSD area.

**Keywords:** Systematization of knowledge, Distributed software development, Taxonomy, Empirical evidence, Empirical studies.



## LISTA DE FIGURAS

FIGURA 1 - ELEMENTOS DA TAXONOMIA PROPOSTA POR [SMI12].	35
FIGURA 2 - PERCENTUAL DE ARTIGOS POR LOCAL.	38
FIGURA 3 - PERCENTUAL DE ARTIGOS POR LOCAL.	39
FIGURA 4 - COMPARAÇÃO DOS RESULTADOS NO PERÍODO DE 2000 A 2007.	40
FIGURA 5 - ETAPAS DO MAPEAMENTO SISTEMÁTICO.	43
FIGURA 6 - ANO DE PUBLICAÇÃO DOS ARTIGOS IDENTIFICADOS.	48
FIGURA 7 - NÚMERO DE ESTUDOS POR DÉCADA.	49
FIGURA 8 - DISTRIBUIÇÃO DOS ARTIGOS POR ÁREA DO SWEBOK.	50
FIGURA 9 - PERCENTUAL DE ARTIGOS POR ÁREA DO SWEBOK.	51
FIGURA 10 - DISTRIBUIÇÃO POR ESTRATÉGIA DE SISTEMATIZAÇÃO.	51
FIGURA 11 - DISTRIBUIÇÃO POR ESTRATÉGIA AO LONGO DOS ANOS.	52
FIGURA 12 - DESENHO DE PESQUISA.	74
FIGURA 13 – TEMPO DE EXPERIÊNCIA COMO PESQUISADOR EM DDS.	84
FIGURA 14 – DISTRIBUIÇÃO GEOGRÁFICA DOS RESPONDENTES DA ACADEMIA.	84
FIGURA 15 - PERCEPÇÃO DA ACADEMIA SOBRE A BUSCA POR ARTIGOS EMPÍRICOS.	88
FIGURA 16 – RESULTADOS DA ACADEMIA (ARTIGO A).	90
FIGURA 17 – RESULTADOS DA ACADEMIA (ARTIGO B).	92
FIGURA 18 – TEMPO DE EXPERIÊNCIA DOS PROFISSIONAIS DA INDÚSTRIA COM DDS.	94
FIGURA 19 - DISTRIBUIÇÃO GEOGRÁFICA DOS RESPONDENTES DA INDÚSTRIA.	95
FIGURA 20 - PERCEPÇÃO DA INDÚSTRIA SOBRE A BUSCA POR ARTIGOS EMPÍRICOS.	97
FIGURA 21 – RESULTADOS DA INDÚSTRIA (ARTIGO A).	99
FIGURA 22 – RESULTADOS DA INDÚSTRIA (ARTIGO B).	101
FIGURA 23 – TAXONOMIA PROPOSTA.	112
FIGURA 24 - PERCENTUAL DE ARTIGOS EMPÍRICOS.	116



## LISTA DE TABELAS

TABELA 1 – PRINCIPAIS DESAFIOS DO DDS. ....	30
TABELA 2 - DESAFIOS DE DDS QUE AINDA NÃO APRESENTAM SOLUÇÕES NA LITERATURA. ....	31
TABELA 3 – LISTA DE LOCAIS, ACRÔNIMOS E NÚMERO DE ARTIGOS DA RSL DE 2010.....	37
TABELA 4 - LISTA DE LOCAIS, ACRÔNIMOS E NÚMERO DE ARTIGOS DA RSL DE 2012. ....	39
TABELA 5 - COMPARAÇÃO DAS REVISÕES SISTEMÁTICAS POR LOCAL DE PUBLICAÇÃO.....	41
TABELA 6 - NÚMERO DE ARTIGOS POR ETAPA.....	47
TABELA 7 - DISTRIBUIÇÃO DOS ARTIGOS POR ÁREA E ESTRATÉGIA DE SISTEMATIZAÇÃO.....	53
TABELA 8 - SISTEMATIZAÇÃO EM PROCESSO DE ENGENHARIA DE SOFTWARE.....	54
TABELA 9 - SISTEMATIZAÇÃO EM DESIGN DE SOFTWARE. ....	60
TABELA 10 - SISTEMATIZAÇÃO EM CONSTRUÇÃO DE SOFTWARE. ....	62
TABELA 11 - SISTEMATIZAÇÃO EM REQUISITOS DE SOFTWARE.....	64
TABELA 12 - SISTEMATIZAÇÃO EM FERRAMENTAS E MÉTODOS DE ES. ....	66
TABELA 13 - SISTEMATIZAÇÃO EM QUALIDADE DE SOFTWARE.....	68
TABELA 14 - PERFIL DOS RESPONDENTES DA <i>SURVEY</i> . ....	81
TABELA 15 – SÍNTESE DAS MOTIVAÇÕES DA ACADEMIA PARA BUSCA DE ESTUDOS EMPÍRICOS. ....	86
TABELA 16 – SÍNTESE DAS EXPETATIVAS DA ACADEMIA. ....	87
TABELA 17 – SÍNTESE DOS DESAFIOS DA ACADEMIA. ....	90
TABELA 18 – SÍNTESE DAS INFORMAÇÕES DE CONTEXTO (ACADEMIA). ....	93
TABELA 19 – SÍNTESE DAS MOTIVAÇÕES DA INDÚSTRIA PARA BUSCA DE ESTUDOS EMPÍRICOS.....	96
TABELA 20 – SÍNTESE DAS EXPETATIVAS DA INDÚSTRIA. ....	97
TABELA 21 – SÍNTESE DOS DESAFIOS DA INDÚSTRIA.....	99
TABELA 22 – SÍNTESE DAS INFORMAÇÕES DE CONTEXTO (INDÚSTRIA). ....	103
TABELA 23 – CONSOLIDAÇÃO DOS RESULTADOS DA <i>SURVEY</i> .....	104
TABELA 24- CATEGORIAS, ELEMENTOS E FONTE DA TAXONOMIA PROPOSTA. ....	113
TABELA 25 - CENÁRIOS DE AVALIAÇÃO DA TAXONOMIA.....	115
TABELA 26 – CONFERÊNCIA, ANO E NÚMERO DE ARTIGOS.....	115
TABELA 27 - LISTA DE ARTIGOS DO CENÁRIO C1. ....	117
TABELA 28 - LISTA DE ARTIGOS DO CENÁRIO C3. ....	118
TABELA 29 - LISTA DE ARTIGOS DO CENÁRIO C4. ....	118
TABELA 30 - ARTIGO COM MÚLTIPLAS CLASSIFICAÇÕES (FONTE DE EVIDÊNCIA EMPÍRICA). ....	119
TABELA 31 - ARTIGO COM MÚLTIPLAS CLASSIFICAÇÕES (LOCALIZAÇÃO). ....	119



## LISTA DE SIGLAS

DDS – Desenvolvimento Distribuído de Software

ES – Engenharia de Software

GSD – *Global Software Development*

GSE – *Global Software Engineering*

ICGSE - *International Conference on Global Software Engineering*

PEP – Plano de Estudo e Pesquisa

PUCRS – Pontifícia Universidade Católica do Rio Grande do Sul

RSL – Revisão Sistemática da Literatura

TI – Tecnologia da Informação



## SUMÁRIO

1	INTRODUÇÃO .....	23
1.1	<b>OBJETIVOS</b> .....	25
1.2	<b>JUSTIFICATIVA</b> .....	26
1.3	<b>ORGANIZAÇÃO DO TRABALHO</b> .....	27
2	REFERENCIAL TEÓRICO .....	28
2.1	<b>DESENVOLVIMENTO DISTRIBUÍDO DE SOFTWARE</b> .....	28
2.1.1	Fatores que levam ao DDS.....	29
2.1.2	Desafios do DDS .....	30
2.2	<b>ESTUDOS SOBRE SISTEMATIZAÇÃO DAS EVIDÊNCIAS EMPÍRICAS EM DDS</b> .....	32
2.2.1	Smite et al. 2010 [SMI10] .....	32
2.2.2	Smite et al. 2012 [SMI12] .....	33
2.2.3	Outros estudos .....	36
2.3	<b>DESAFIOS DA SISTEMATIZAÇÃO DAS EVIDÊNCIAS EMPÍRICAS EM DDS</b> .....	37
2.3.1	Análise do processo de RSL em DDS .....	37
2.3.2	Comparação dos resultados de cada RSL.....	39
2.3.3	Discussão.....	41
2.4	<b>SISTEMATIZAÇÃO DO CONHECIMENTO EM ENGENHARIA DE SOFTWARE</b> .....	42
2.4.1	Protocolo do mapeamento sistemático .....	43
2.4.2	Identificação da literatura relevante .....	44
2.4.3	Seleção dos estudos .....	45
2.4.4	Avaliação da qualidade.....	46
2.4.5	Extração dos dados.....	46
2.4.6	Execução.....	47
2.4.7	Análise dos resultados.....	47
2.4.7.1	Análise quantitativa .....	48
2.4.7.2	Análise qualitativa .....	52
2.4.8	Conclusões do mapeamento sistemático.....	69
2.4.9	Limitações do mapeamento sistemático .....	72
3	METODOLOGIA DE PESQUISA.....	73
3.1	<b>DESENHO DE PESQUISA</b> .....	74
3.2	<b>ASPECTOS METODOLÓGICOS</b> .....	77
4	SURVEY.....	78
4.1	<b>BASE METODOLÓGICA</b> .....	78
4.1.1	Seleção das organizações e unidade de análise .....	79
4.1.2	Fonte de dados e seleção de participantes.....	79

4.1.3	Análise de Dados .....	79
<b>4.2</b>	<b>FASES E OPERACIONALIZAÇÃO DA SURVEY</b> .....	<b>80</b>
4.2.1	Perfil da população .....	81
4.2.2	Roteiro de entrevista .....	81
<b>4.3</b>	<b>RESULTADOS DA SURVEY</b> .....	<b>83</b>
4.3.1	Resultados da academia .....	83
4.3.1.1	Caracterização dos respondentes da academia .....	83
4.3.1.2	Elementos de análise da academia .....	84
4.3.2	Resultados da indústria .....	94
4.3.2.1	Caracterização dos respondentes da indústria .....	94
4.3.2.2	Elementos de análise da indústria .....	95
<b>4.4</b>	<b>CONSOLIDAÇÃO DOS RESULTADOS DA SURVEY</b> .....	<b>103</b>
<b>4.5</b>	<b>LIÇÕES PARA O ESTUDO</b> .....	<b>105</b>
5	TAXONOMIA .....	108
<b>5.1</b>	<b>CATEGORIAS E ELEMENTOS DA TAXONOMIA PROPOSTA</b> .....	<b>109</b>
<b>5.2</b>	<b>PROCESSO DE AVALIAÇÃO DA TAXONOMIA</b> .....	<b>114</b>
5.2.1	Cenários .....	114
5.2.2	Seleção dos estudos .....	115
5.2.3	Extração dos dados .....	115
<b>5.3</b>	<b>RESULTADOS DA AVALIAÇÃO</b> .....	<b>116</b>
5.3.1	Resultados por cenário .....	116
5.3.2	Estudos com múltiplas classificações .....	118
6	CONSIDERAÇÕES FINAIS .....	120
<b>6.1</b>	<b>CONTRIBUIÇÕES DA PESQUISA</b> .....	<b>121</b>
<b>6.2</b>	<b>LIMITAÇÕES DA PESQUISA</b> .....	<b>121</b>
<b>6.3</b>	<b>PESQUISAS FUTURAS</b> .....	<b>122</b>
	REFERÊNCIAS BIBLIOGRÁFICAS .....	123
	APÊNDICE A .....	129
	APÊNDICE B .....	134
	APÊNDICE C .....	137
	APÊNDICE D .....	148

# 1 INTRODUÇÃO

Ao longo dos últimos 15 anos, a comunidade de Engenharia de Software (ES) testemunhou o desenvolvimento de uma área de pesquisa relativamente recente, chamada por alguns como Desenvolvimento Distribuído de Software (DDS) e por outros como Engenharia de Software Global [PRI13], [BEE13], [CAR99], [HER07]. Hoje em dia, as empresas estão cada vez mais distribuindo seus processos de desenvolvimento de software em diversas regiões no mundo, aproveitando-se de melhorias tecnológicas e regimes de comércio global [PRI13]. A partir deste novo cenário, o DDS<sup>1</sup> está ganhando cada vez mais espaço entre as grandes organizações, onde se observou um grande investimento na conversão de mercados locais em globais, criando novas formas de competição e colaboração para o desenvolvimento de software [HER01].

O DDS evoluiu recentemente, resultando em um grande número de pesquisas e estudos da indústria [SMI12]. Muitas empresas aderiram ao DDS para se beneficiar do desenvolvimento mais barato, mais rápido e melhor de sistemas de software, produtos e serviços [SMI10]. A principal vantagem reside em uma maior disponibilidade de recursos humanos em zonas descentralizadas de menor custo. Há, contudo, algumas desvantagens que são causadas pela distância que separa as equipes de desenvolvimento, assim como, a coordenação e a comunicação tornam-se mais difícil, pois os componentes do software são provenientes de locais diferentes, afetando a organização do projeto, controle de projetos, e qualidade do produto [JIM09].

Ao longo dos últimos anos surgiram diversas estratégias para operacionalizar o DDS, devido à variedade de possibilidades existentes. Além disso, considerada sua recente aplicação e grande abrangência, o DDS possui uma diversidade de conceitos que auxiliam na sua caracterização [AUD07]. O DDS é um fenômeno que vem crescendo desde a última década e tem sido caracterizado pela colaboração e cooperação entre departamentos de organizações e pela criação de grupos de desenvolvedores que trabalham em conjunto, localizados em cidades ou países diferentes (MEY06).

Segundo Smite et al. [SMI12], muitas organizações nos dias de hoje competem através da utilização dos benefícios oferecidos pelo DDS e estratégias de *sourcing* são, portanto, discutidas com mais frequência. Uma vez que existem muitas variações de atributos associados com projetos de DDS, uma grande quantidade de novos termos foi

---

<sup>1</sup> Existem diversas formas de caracterizar o desenvolvimento de software com equipes dispersas, tais como DDS (Desenvolvimento Distribuído de Software) [LAY00], GSD (*Global Software Development*) [CAR99], *Multi-site development* [KAR98] e GSE (*Global Software Engineering*) ou Engenharia de Software Global [HER07]. Neste estudo, adota-se o termo DDS.

introduzida. A diversidade do jargão *sourcing*, por exemplo, tem causado dificuldades em determinar qual o termo devemos usar em qual situação, e assim causando novos obstáculos para buscar e encontrar pesquisa relevante durante, por exemplo, uma revisão sistemática da literatura. A incapacidade de julgar a aplicabilidade da pesquisa em um contexto industrial é outra implicação importante na transferência de conhecimento para a prática.

Smite et al. [SMI12] complementa que a terminologia usada para o DDS não é padronizada. Os cenários de DDS são diversos e o que se aplica em um contexto talvez não se aplique diretamente em outro. Dessa forma é necessário entender como os resultados empíricos em DDS deveriam ser relatados para ser útil para profissionais na indústria e pesquisadores. A pesquisa empírica na área de DDS poderia ser melhorada se houvesse uma maneira sistemática para identificar as diferentes dimensões e características para o qual se aplica a evidência empírica.

Holliday [HOL06] descreve em seu livro que a sistematização é aquela interpretação crítica de uma ou várias experiências que, a partir de seu ordenamento e reconstrução, descobre ou explicita a lógica do processo vivido, os fatores que intervieram no dito processo, como se relacionaram entre si e porque o fizeram desse modo. O autor ainda acrescenta outras características importantes da sistematização:

- A sistematização de uma experiência produz um novo conhecimento, um primeiro nível de conceitualização a partir da prática concreta que, uma vez que possibilita sua compreensão, leva a transcendê-la, a ir mais além dela mesma. Nesse sentido, permite-nos abstrair o que estamos fazendo em cada caso particular e encontrar um terreno fértil onde a generalização é possível;
- A sistematização ordena conhecimentos e percepções dispersas que surgiram no transcorrer da experiência. Assim, explicita intuições, intenções e vivências acumuladas ao longo do processo. Ao sistematizar, as pessoas recuperam de maneira ordenada o que já sabem sobre sua experiência e descobrem o que ainda não sabem.

O autor também destaca de forma resumida as múltiplas possibilidades e utilidades da sistematização do seguinte modo:

- Ter uma compreensão mais profunda das experiências que realizamos visando melhorar nossa própria prática;

- Compartilhar com outras práticas semelhantes os ensinamentos surgidos com a experiência;
- Conduzir à reflexão teórica (e em geral à construção de teoria) os conhecimentos surgidos de práticas sociais concretas.

Holliday [HOL06] descreve que o exercício de sistematização é um exercício claramente teórico, ou seja, é um esforço rigoroso que formula categorias, classifica e ordena elementos empíricos, faz análise e síntese, indução e dedução e obtém conclusões e as formula como pautas para sua verificação prática. A sistematização relaciona os processos imediatos com seus contextos, confronta o fazer prático com os pressupostos teóricos que o inspiram. Assim, o processo de sistematização se sustenta em uma fundamentação teórica e filosófica sobre o conhecimento e sobre a realidade histórico-social.

O tema desta pesquisa está neste contexto exploratório e tem por objetivo propor uma forma de sistematização das evidências empíricas em DDS, visando contribuir com o desenvolvimento de novas pesquisas nesta área. A evidência empírica é uma fonte de conhecimento adquirido por meio da observação ou experimentação [WIK14]. Neste estudo, busca-se explorar as formas de sistematização do conhecimento na área de Engenharia de Software (ES), mapeando os desafios e os resultados encontrados na literatura e da mesma forma, busca-se compreender que tipo de informação de contexto é relevante para pesquisadores e profissionais da indústria, em estudos com evidências empíricas em DDS. Desta forma, a questão de pesquisa que norteou este estudo foi:

*Como podemos sistematizar as evidências empíricas geradas a partir das pesquisas científicas conduzidas na área de Desenvolvimento Distribuído de Software?*

## **1.1 OBJETIVOS**

O objetivo geral desta dissertação de mestrado é propor uma forma de sistematizar as evidências empíricas geradas a partir das pesquisas científicas conduzidas na área de Desenvolvimento Distribuído de Software. O uso de ontologias [CHA99], como forma de sistematização do conhecimento, não faz parte do escopo deste trabalho. Decidiu-se durante o Plano de Estudo e Pesquisa que o foco dessa pesquisa seria propor uma taxonomia, conforme descrito nos objetivos específicos a seguir.

Visando contemplar o objetivo geral proposto, os seguintes objetivos específicos foram definidos:

- Identificar estudos sobre sistematização do conhecimento em Engenharia de Software, através da realização de uma revisão sistemática da literatura, do tipo mapeamento sistemático;
- Mapear as informações de contexto, geradas a partir de pesquisas científicas conduzidas na área de DDS, que são relevantes para indústria e academia, através de um estudo exploratório e empírico;
- Propor uma taxonomia para sistematização de evidências empíricas em DDS;
- Avaliar a taxonomia proposta através da classificação de estudos empíricos em DDS.

## 1.2 JUSTIFICATIVA

Sistematizar evidências empíricas em DDS não é uma tarefa trivial e possui alguns desafios, conforme relatado em estudos recentes da área [SMI12], [SMI10], [PRI10] e [JIM09]:

- O DDS ainda é uma área imatura, tornando a sistematização das evidências empíricas uma tarefa complexa;
- Muitas vezes o contexto de um estudo empírico em DDS não é descrito de forma clara, o que faz com que seja difícil para pesquisadores e profissionais identificarem casos que podem ser de interesse para eles;
- A terminologia aplicada ao DDS não é consistente e padronizada, uma vez que existem diversos estudos que utilizam os mesmos termos de forma diferente. Os pesquisadores têm enfrentado dificuldades em buscar e encontrar pesquisa relevante, uma vez que as palavras-chave usadas neste campo são muito diversas.

No artigo [SMI10] os autores abordam a capacidade dos leitores e pesquisadores em encontrar estudos empíricos em DDS. Eles relatam que mesmo usando os termos mais populares relacionados com a área em uma *string* de busca, o resultado revelou 387 estudos, mas apenas 59 artigos foram relevantes, ou seja, apenas 15% dos estudos encontrados eram qualificados para a revisão sistemática. Para permitir que os profissionais da indústria e os pesquisadores entendam o estado da arte em DDS, deve-se relatar os estudos de forma mais consistente. Este é um pré-requisito para realmente identificar o que se sabe e o que não se sabe sobre engenharia de software global. Uma

revisão sistemática da literatura precisa ser capaz de deduzir os dados de um grande número de artigos sem ambiguidade ou incertezas [SMI08]. É neste contexto que a pesquisa de como sistematizar estudos empíricos em DDS mostra-se relevante.

### 1.3 ORGANIZAÇÃO DO TRABALHO

Essa dissertação está dividida em seis capítulos, sendo indicada a seguir a estrutura do presente trabalho:

No Capítulo 2 apresenta-se o referencial teórico desta pesquisa, envolvendo os principais conceitos e implicações das áreas de estudo: desenvolvimento distribuído de software, sistematização das evidências empíricas em DDS e um estudo de mapeamento sistemático sobre a sistematização do conhecimento em ES. O Capítulo 3 tratará da metodologia de pesquisa utilizada, descrevendo as etapas do estudo, justificando a escolha e uso dos métodos.

No Capítulo 4 apresenta-se a *survey*, descrevendo a metodologia, as fases e os resultados obtidos. A *survey* com especialistas identifica quais evidências empíricas e informações de contexto dos estudos em DDS são relevantes para profissionais e pesquisadores desta área.

A taxonomia proposta para sistematização das evidências empíricas em DDS é apresentada no Capítulo 5, como consequência do processo de pesquisa seguido, apoiado nos resultados do mapeamento sistemático e da *survey* realizada. No Capítulo 6 apresentam-se as considerações finais sobre o tema e enfocam-se os aspectos relacionados às contribuições e limitações deste estudo. Conclui-se o trabalho destacando rumos para futuras pesquisas sobre o tema.

## 2 REFERENCIAL TEÓRICO

O referencial teórico constitui uma importante etapa da pesquisa, contendo os principais conceitos e teorias das áreas pesquisadas. Dessa forma, na seção 2.1, apresenta-se a área de Desenvolvimento Distribuído de Software (DDS), contemplando as definições, características, fatores que levaram a sua utilização e os desafios relacionados. A Seção 2.2 trata sobre estudos que discutem a sistematização das evidências empíricas em DDS e a seção 2.3 descreve uma análise sobre o processo de revisão sistemática utilizado nos estudos Smite et al. 2010 e 2012 [SMI10 e SMI12]. Por fim, na seção 2.4, apresentam-se os principais desafios e soluções sobre sistematização do conhecimento em Engenharia de Software (ES), através de um estudo de mapeamento sistemático.

### 2.1 DESENVOLVIMENTO DISTRIBUÍDO DE SOFTWARE

Nos últimos anos, pode-se perceber um grande avanço em direção à globalização dos negócios. Empresas de desenvolvimento de software estão distribuindo seus processos em todo o mundo, a fim de atender as necessidades do mercado globalizado [MAR12]. Na área de Engenharia de Software (ES), mercados nacionais têm se transformado em mercados globais, criando novas formas de competição que vão além das fronteiras dos países. Hoje em dia, mais projetos estão sendo desenvolvidos em ambientes geograficamente distribuídos, caracterizando o Desenvolvimento Distribuído de Software (DDS). Projetos de desenvolvimento de software distribuído impõem novos padrões de cooperação entre as entidades organizacionais e atores humanos, que atravessam fronteiras culturais, físicas, temporais e geográficas [PRI13].

O DDS é caracterizado sempre que um ou mais recursos humanos envolvidos no projeto estiverem fisicamente distantes dos demais [AUD07]. Pode-se dizer também que uma equipe global de desenvolvimento de software está tipicamente em países diferentes, porém colaborando em um mesmo projeto de qualquer natureza (criação ou manutenção de software) [LAN08].

Carmel [CAR99] conceitua o DDS como um modelo de desenvolvimento onde os envolvidos em um determinado projeto estão dispersos, sendo este diferente do co-localizado quanto à dispersão geográfica, temporal e diferenças culturais. Audy e Prikladnicki [AUD07] indicam que essas três características acabam criando diferentes sensações de distância, que por sua vez se multiplicam em diversas dificuldades de

coordenação do trabalho necessário para desenvolver produtos de software, refletindo em diversos fatores. Herbsleb e Moitra [HER01] indicam que entre estes fatores destacam as questões:

- Estratégicas, referentes à decisão de desenvolver ou não um projeto de software de forma distribuída, tomando por base a análise de risco e custo-benefício;
- Culturais, que envolvem valores e princípios entre as equipes distribuídas;
- Técnicas, de acordo com fatores relativos à infraestrutura tecnológica e ao conhecimento técnico, tais como rede de comunicação de dados, plataformas de hardware, processo de desenvolvimento, etc.;
- De gestão de conhecimento, sendo fatores relativos à criação, armazenamento, processamento e compartilhamento de informações nos projetos distribuídos.

### 2.1.1 Fatores que levam ao DDS

As atividades do desenvolvimento de software realizadas de forma distribuída vêm aumentando nos últimos anos, em decorrência de um novo cenário apresentado a partir da globalização dos negócios. Pesquisadores têm estudado DDS por mais de uma década e os principais fatores que contribuíram para o crescimento do DDS nos últimos anos foram [PRI13], [BEE13], [PRI03], [CAR99]:

- A necessidade de recursos globais para serem utilizados a qualquer hora;
- Incentivos fiscais para o investimento em pesquisas de informática;
- Disponibilidade de mão-de-obra especializada e de custos reduzidos em países em desenvolvimento;
- Vantagem de estar perto do mercado local, incluindo o conhecimento dos clientes e as condições locais para explorar as oportunidades de mercado;
- Rápida formação de organizações e equipes virtuais para explorar as oportunidades locais;
- Grande pressão para o desenvolvimento *time-to-market* (velocidade no trabalho, tempo entre a concepção e a comercialização do produto) utilizando as vantagens do fuso horário diferente, no desenvolvimento conhecido como *follow-the-sun* (24 horas contínuas, cotando com equipes fisicamente distribuídas); e
- Necessidade de integrar recursos resultantes de aquisições e fusões organizacionais.

### 2.1.2 Desafios do DDS

Assim como ocorre no desenvolvimento co-localizado de software, o distribuído também possui vários problemas e desafios inerentes ao processo. Neste sentido, Audy e Prikladnicki [AUD07] dividiram em cinco categorias (pessoas, processo, tecnologia, gestão e comunicação) os principais desafios verificados no DDS, sendo exibidos na Tabela 1.

Tabela 1 – Principais desafios do DDS.

Categories	Desafios
Pessoas	Confiança, Conflitos, Diferenças culturais, Ensino de DDS, Espírito de equipe, Formação de equipes e grupos, Liderança e Tamanho de equipe.
Processo	Arquitetura de software, Engenharia de requisitos, Gerência de configuração e Processo de desenvolvimento.
Tecnologia	Tecnologias de colaboração e Telecomunicações.
Gestão	Coordenação, Controle e interdependência, Gestão de portfólio de projetos, Gerência de risco, Legislação, Modelos de negócio e Seleção e alocação de projetos.
Comunicação	<i>Awareness</i> (percepção), Contexto, Dispersão geográfica e temporal, Estilo e formas de comunicação e Fusos horários.

Fonte: [AUD07]

Os desafios relacionados a pessoas dizem respeito às características que afetam diretamente os recursos humanos envolvidos no desenvolvimento de software. A liderança também é considerada um importante desafio para o DDS. Líderes em algumas culturas costumam tratar as equipes de forma participativa e democrática, enquanto outros tratam a liderança de forma mais hierárquica, tomando decisões sem consultar seus subordinados.

Os desafios da categoria processo referem-se à forma como o projeto é desenvolvido. É importante, por exemplo, ter uma arquitetura do software que seja modularizada, pois o DDS exige uma colaboração constante e com uma menor interdependência entre os locais de desenvolvimento.

A telecomunicação e as tecnologias de colaboração englobam os desafios referentes à tecnologia no DDS. A partir da distância física entre as equipes, os recursos tecnológicos se fazem necessários para facilitar a comunicação e o trabalho colaborativo entre os membros da equipe. As tecnologias de colaboração (*groupware*) compreendem ferramentas síncronas e assíncronas e que podem ser utilizadas de pelas equipes de desenvolvimento, estejam elas localizadas de forma dispersa ou local.

Dentro dos níveis de gestão organizacional (estratégica, tática e operacional) os desafios de gestão estão relacionados ao gerenciamento de projetos, de risco, legislação,

modelos de negócio, seleção e alocação de projetos. A coordenação em ambientes distribuídos é considerado um importante desafio à gestão, pelos problemas de cultura, idioma e tecnologia serem ampliados devido a distância.

Os desafios de comunicação contribuem para o relacionamento e a gestão da infraestrutura das equipes e *stakeholders* envolvidos em projetos de DDS. A percepção (*awareness*), contexto, dispersão geográfica e temporal, estilo e formas de comunicação e os fusos horários são alguns dos desafios relacionados à comunicação. É importante que as equipes conheçam e saibam como tratar os diferentes fusos horários nas diversas localidades.

Gomes e Marczak [GOM12] executaram uma revisão sistemática com o objetivo de identificar quais desafios e problemas de DDS possuem soluções publicadas na literatura. Este estudo visou consolidar o corpo de conhecimento sobre o tema, a fim de facilitar a identificação de futuras pesquisas e melhorias necessárias por acadêmicos e profissionais da indústria. Na Tabela 2 listamos os desafios encontrados, por categoria, que ainda não apresentam uma solução publicada na literatura.

Tabela 2 - Desafios de DDS que ainda não apresentam soluções na literatura.

<b>Categorias</b>	<b>Desafios sem soluções publicadas na literatura</b>
Pessoas	<ul style="list-style-type: none"> <li>• Medo de perder o emprego;</li> <li>• Falta de experiência de trabalho em equipes distribuídas.</li> </ul>
Processo	<ul style="list-style-type: none"> <li>• Dificuldade em definir gerenciamento de configuração e processo de controle de versionamento eficazes;</li> <li>• Dificuldade em estabelecer o entendimento sobre as práticas de trabalho diversificadas de diferentes sites;</li> <li>• Dificuldade na definição de um processo de negócio que atenda as exigências dos diversos sites;</li> <li>• Dificuldade em definir e institucionalizar um processo unificado que cubra todo o ciclo de vida.</li> </ul>
Tecnologia	<ul style="list-style-type: none"> <li>• Infraestrutura, ferramentas e técnicas limitadas para o apoio ao DDS;</li> <li>• Tecnologias limitadas de telecomunicações.</li> </ul>
Gestão	<ul style="list-style-type: none"> <li>• Dificuldade em gerenciar a falta de a cooperação das partes interessadas;</li> <li>• Dificuldade em lidar com diversas leis e políticas;</li> <li>• Dificuldade no planejamento de estimativa de esforço acurada;</li> <li>• Tempo limitado para colaboração com colegas remotos.</li> </ul>
Comunicação	<ul style="list-style-type: none"> <li>• Dificuldade em estabelecer um entendimento comum entre os membros que utilizam vocabulários e terminologias diferentes;</li> <li>• Poucas oportunidades para comunicação informal e estabelecimento de relações interpessoais;</li> <li>• Formas eficazes para estabelecer a comunicação são limitadas;</li> <li>• Dificuldade para troca de informações e fluxo de comunicação eficiente.</li> </ul>

Fonte [GOM12]

Categorias como processo, gestão e comunicação apresentam diversos desafios sem uma solução publicada na literatura, gerando, portanto, uma demanda por novos estudos abordando esses temas.

## 2.2 ESTUDOS SOBRE SISTEMATIZAÇÃO DAS EVIDÊNCIAS EMPÍRICAS EM DDS

Nesta seção apresentamos estudos que discutem a sistematização das evidências empíricas em DDS. Dessa forma, nas seções 2.2.1 e 2.2.2, apresentam-se os dois principais estudos relacionados com esta pesquisa e na seção 2.2.3, descrevem-se sucintamente outros estudos relacionados com o tema.

### 2.2.1 Smite et al. 2010 [SMI10]

Este foi o primeiro estudo a realizar uma revisão sistemática da área de DDS com foco em estudos empíricos. Os principais objetivos dos autores foram sistematizar os desafios empiricamente documentados e procurar evidências empíricas sobre soluções para os desafios de DDS relatados na literatura. Para sistematizar os desafios, os autores criaram um esquema de classificação para extrair os dados dos artigos de forma categorizada. O esquema inclui a captura de dados sobre a relevância, contexto empírico, contexto de DDS, contexto do estudo e uma avaliação qualitativa. A relevância diz respeito ao campo de estudo (DDS ou ES), o contexto empírico especifica a metodologia utilizada no estudo, o contexto de DDS classifica o estudo conforme as características de DDS, tais como, distribuição geográfica, modo de colaboração, número de empresas envolvidas, etc. Por fim o contexto de estudo e da avaliação qualitativa descrevem detalhes do estudo, tais como, domínio da aplicação, foco do estudo, caso de sucesso ou fracasso, recomendações, entre outros.

Os autores observaram que a maioria dos estudos representa problemas orientados a relatórios com foco em diferentes aspectos da gestão de DDS, em vez de uma análise em profundidade de soluções, por exemplo, em termos de práticas úteis ou técnicas. As empresas ainda estão impulsionadas por estratégias de redução de custos e, ao mesmo tempo, as recomendações mais frequentemente discutidas indicam uma necessidade de investimentos em viagens e socialização. Assim, ao mesmo tempo em que o desenvolvimento se torna global, há uma ambição para minimizar a separação geográfica, temporal e cultural.

O artigo também relata que a maioria das pesquisas empíricas parece ser de natureza exploratória e se concentram em identificar os problemas causados pela distribuição e possíveis soluções são discutidas em termos muito gerais. No entanto, ainda não há um padrão ou receita de sucesso do DDS. Além disso, as lições aprendidas muitas vezes focam na estruturação de fatores de sucesso, mas dão pouca orientação sobre como realmente alcançar o sucesso. Devido aos diversos cenários ativados por

diferentes formas de trabalho colaborativo global, lições aprendidas em um contexto não se aplicam diretamente em outro contexto [SMI08]. Isto significa que a aprendizagem da literatura existente pode exigir um esforço considerável para os profissionais entenderem a aplicabilidade dos resultados oferecidos, colocando exigências crescentes sobre os pesquisadores para que explicitamente e claramente descrevam o contexto e as circunstâncias ao relatar estudos empíricos.

### 2.2.2 Smite et al. 2012 [SMI12]

Este estudo apresenta uma proposta de taxonomia para área de DDS categorizando termos previamente selecionados com base no relacionamento de generalização-especialização e mostram como a taxonomia pode ser usada para classificar e mapear o conhecimento existente em DDS. Os autores citam que a diversidade no termo *outsourcing*, por exemplo, tem causado dificuldades em determinar qual termo usar em que situação e dessa forma acabam criando novos obstáculos para buscar e encontrar estudos relevantes durante uma revisão sistemática da literatura. Assim, a necessidade de criar uma terminologia precisa, com definições para as diferentes situações de *sourcing* global surge como um caminho para a comunidade desenvolver novos trabalhos com base em outros estudos e, portanto, progredir mais rapidamente.

O objetivo dessa pesquisa foi investigar o estado de utilização das terminologias que caracterizam estratégias de *sourcing* em DDS em relação à consistência na ortografia e significado, e como resultado, um conjunto de termos foi analisado por diversos pesquisadores experientes na área e posteriormente categorizados na forma de uma taxonomia. Segundo os autores, uma taxonomia acrescenta a uma terminologia, descrevendo as relações entre os termos. A taxonomia descreve as relações de generalização para especialização, o que ajuda a ligar termos de uma área ou de um tema. Assim, uma taxonomia é uma classificação hierárquica de um tema ou área. O público alvo deste artigo visa pesquisadores, que publicam ou sintetizam estudos empíricos em DDS e profissionais, que estão interessados em casos empíricos publicados.

Este estudo concentrou-se em estratégias de *sourcing* com uma combinação de localização geográfica e estrutura de relacionamento entre empresas. Os autores propõem a construção de uma taxonomia para identificar semelhanças e diferenças entre as estratégias de *sourcing* e argumentam que com base em uma taxonomia, novas

observações ou estudos de caso poderiam ser comparados com a classificação existente. Embora exista uma grande variedade de termos na área de DDS, este artigo visou à compreensão dos termos mais populares, independentemente se os termos eram definidos pelos autores ou não, e se existia um acordo comum no uso dos termos. A taxonomia proposta é dividida em cinco níveis.

1. DDS: o ponto de partida é o *sourcing*, ou seja, alguma forma de desenvolvimento externo, que do ponto de vista dos responsáveis, pode ser um projeto, produto, sistema ou serviço.
2. Localização: o fornecimento pode ser *onshore*, ou seja, dentro do mesmo país, ou *offshore* que é em outro país.
3. Entidade Legal: o desenvolvimento será realizado dentro da mesma empresa ou com outra empresa, por exemplo, por subcontratação. *Insourcing*, o desenvolvimento ocorre em outro site da mesma empresa e *outsourcing* é com uma empresa diferente.
4. Distância geográfica: Define uma subclassificação quando a localização for *onshore* ou *offshore*. Quando for *onshore*, pode ser “próxima”, onde nenhum voo é necessário e conseqüentemente o número de reuniões presenciais é maior, ou “distante”, onde um voo é necessário e isso implica em acréscimo de custo e tempo. Para *offshore*, a classificação pode ser “perto”, onde o tempo de voo é menor que duas horas ou “longe”, no qual o tempo de voo é maior que duas horas.
5. Diferença temporal: Podemos ter diferença de tempo tanto em *onshoring* como em *offshoring*. Portanto para situações *onshore*, podemos classificar como “similar” quando a diferença é igual ou menor que uma hora ou “diferente”, quando for maior que uma hora. Para o contexto *offshore*, podemos classificar como “pequeno” quando a diferença de tempo for menor ou igual a quatro horas ou “grande”, quando a diferença é maior que quatro horas.

Conforme as definições acima, os autores criaram a taxonomia apresentada na Figura 1. No total, quatro combinações são impossíveis e representam os casos em que a distância geográfica é “próxima” no caso *onshoring* ou “perto” no caso *offshoring* e são combinadas com as classificações de diferença de tempo “diferente” e “grande” respectivamente. Por exemplo, um voo é necessário se a diferença de tempo é maior que uma hora, e, portanto, a combinação “próxima” e “diferente” é inviável. Um raciocínio semelhante pode ser feito por combinação “perto” e “grande”.

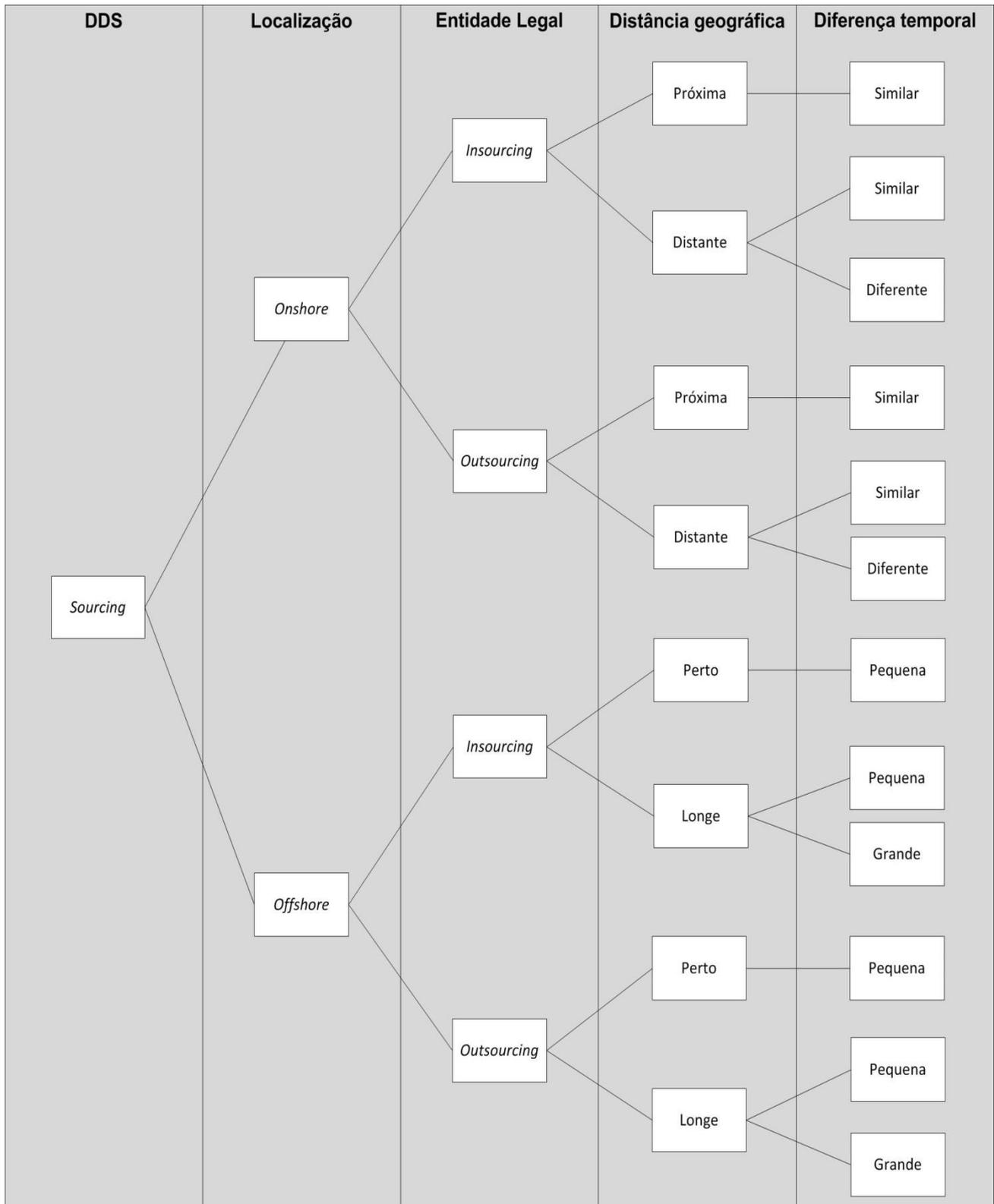


Figura 1 - Elementos da taxonomia proposta por [SMI12].  
Fonte [SMI12]

Smite et al. 2012 [SMI12] descreve que a taxonomia pode ajudar tanto pesquisadores como profissionais da indústria, e facilitar a compreensão dos artigos que documentam descobertas e experiências em DDS. Os autores destacam que a taxonomia

tem dois usos primários. Primeiro, fornece uma base para os pesquisadores classificarem seus próprios estudos e estudos relacionados. Em segundo lugar, uma vez que os estudos são classificados de acordo com a taxonomia, é possível identificar um conjunto de estudos que descrevem uma situação particular. Os pesquisadores podem usar a taxonomia para sintetizar o conhecimento existente, identificar lacunas na literatura e também encontrar trabalhos relacionados, assim como, os profissionais podem usá-la para encontrar evidências de uma pergunta específica em um contexto que é semelhante ao seu. Os autores ainda complementam que o uso de uma terminologia comum e uma taxonomia pode simplificar a comparação entre os estudos em diferentes contextos.

A principal motivação deste estudo está no fato de podermos usar a terminologia e taxonomia propostas em futuras pesquisas sobre DDS. Os autores mencionam que as revisões de literatura em um domínio específico podem se tornar difíceis se os termos não são bem definidos ou padronizados. Este é o caso de novas áreas de pesquisa, tais como DDS. Portanto uma terminologia comum e uma taxonomia em novos campos de pesquisa poderiam tornar mais fáceis o desenvolvimento de novos estudos.

### 2.2.3 Outros estudos

Além dos dois estudos mencionados anteriormente existem outros estudos importantes na área, que serão descritos abaixo de forma resumida.

Prikladnicki e Audy [PRI10] distinguiram entre *offshore outsourcing* (quando o desenvolvimento é movido para uma terceira parte externa em outro país), *internal offshoring* (quando o desenvolvimento é movido para uma divisão de uma empresa específica estabelecida em outro país) e *offshoring* (usado como um termo genérico, quando a relação é desconhecida).

Hossain et al. [HOS09] classificaram modos de colaboração em DDS em duas categorias: inter-organizacionais (entre duas empresas) e intra-organizacional (dentro da mesma empresa).

Gumm [GUM06] propôs uma taxonomia multidimensional usada para descrever como pessoas ou artefatos são distribuídos em um projeto de desenvolvimento de software. As dimensões são: física, organizacional, temporal e a distribuição entre grupos de *stakeholder* medidos na escala alta-média-baixa. No entanto, o trabalho é mais explicativo e se baseia nas dependências causais entre os riscos de distribuição e os mecanismos necessários para fazer lidar com eles ao invés de proporcionar regras concretas que podem ser usados para instanciar a taxonomia ou classificar os casos.

## 2.3 DESAFIOS DA SISTEMATIZAÇÃO DAS EVIDÊNCIAS EMPÍRICAS EM DDS

Esta seção apresenta uma análise crítica sobre os dois estudos que iniciaram uma discussão sobre a sistematização de evidências empíricas em DDS. Dessa forma, na seção 2.3.1, apresenta-se uma análise sobre o processo de revisão sistemática utilizado nos estudos Smite et al. 2010 e 2012 [SMI10 e SMI12]. Na seção 2.3.2 apresentamos uma análise comparativa do resultado das duas revisões sistemáticas e finalmente, na seção 2.3.3, apresenta-se uma discussão sobre os resultados dessa análise.

Como resultado dessa análise crítica, um artigo foi publicado no VII Workshop de Desenvolvimento Distribuído de Software – WDDS 2013 [TEC13].

### 2.3.1 Análise do processo de RSL em DDS

No artigo Smite et al. 2010 [SMI10], os autores utilizaram uma *string* de busca em sete bases de dados para encontrar evidências empíricas ou experiências na área de DDS. O período de publicação considerado foi de 2000 até 2007. Os trabalhos publicados antes de 2000 não foram incluídos na pesquisa e segundo os autores, a principal motivação deve-se ao fato que DDS, com o efeito da globalização, é reconhecido como uma tendência do século 21 [FRI05], e que os estudos realizados após o ano de 2000 são mais relevantes do que os estudos de 10 anos atrás ou mais.

Tabela 3 – Lista de locais, acrônimos e número de artigos da RSL de 2010.

Smite et al. 2010		
Local	Acrônimo	Número artigos
<i>Empirical Software Engineering</i>	ESEJ	2
<i>International Conference on Software Engineering</i>	ICSE	2
<i>IEEE Software</i>	IEEE Software	3
<i>International Workshop on Global Software Development for the Practitioner</i>	GSDPW	3
<i>Information and Software Technology</i>	IST	4
<i>International Conference on Product-Focused Software Process Improvement</i>	PROFES	4
<i>Software Process Improvement and Practice</i>	SPIP	5
<i>International Conference on Global Software Engineering</i>	ICGSE	12
Outros locais com apenas um artigo selecionado	Outros	24
	<b>Total geral</b>	<b>59</b>

Após a execução da *string* de busca, o resultado revelou 387 estudos, mas após um processo rigoroso de seleção apenas 59 artigos foram considerados relevantes pelos autores. O critério de inclusão definido pelos autores pressupõe que o estudo seja claramente empírico e que esteja inserido em um ambiente distribuído, e que o artigo tenha um componente claro de ES. Um componente de ES pode incluir um processo de desenvolvimento completo, ou um estudo de um método específico, técnica ou ferramenta em um contexto DDS. Os locais de publicação que apresentaram mais de um artigo estão listados na Tabela 3. Os locais com apenas um artigo publicado foram agrupados no item “outros” e representam um total de 24 artigos.

Observando o gráfico da Figura 2, no qual representa o número de artigos por local (revista, conferencia ou workshop), nota-se que 20% dos artigos foram publicados na *International Conference on Global Software Engineering (ICGSE)*. Isso é compreensível, uma vez que a ICGSE é dedicada ao tema. As revistas *Software Process Improvement and Practice (SPIP)* e *Information and Software Technology (IST)* representaram 9% e 7% respectivamente.

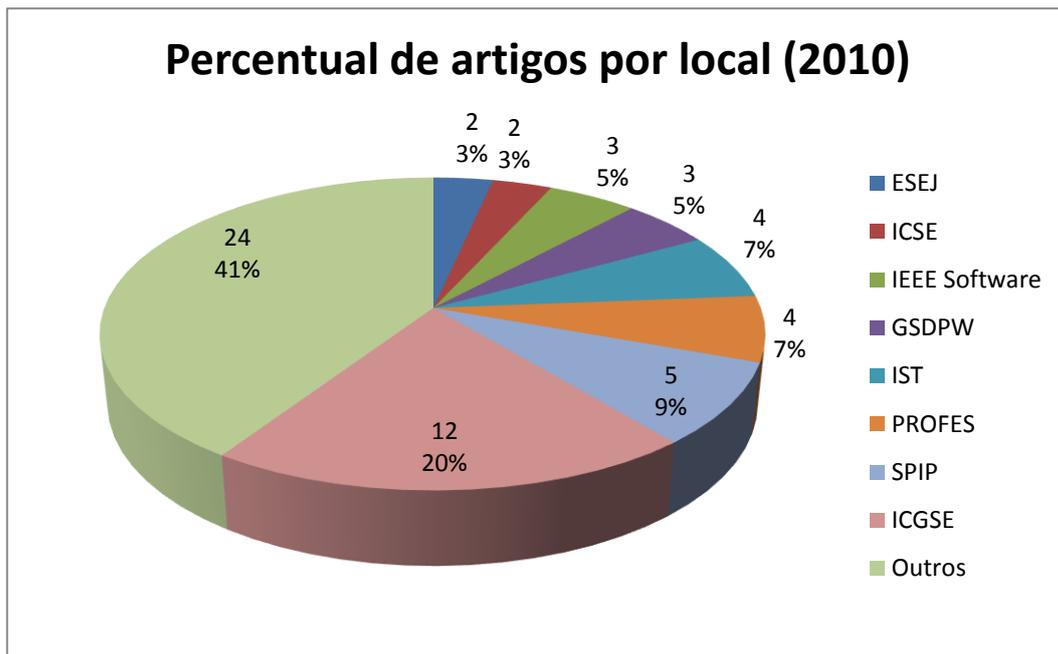


Figura 2 - Percentual de artigos por local.

No estudo Smite et al. 2012 [SMI12], a fim de investigar o estado de utilização das terminologias em DDS, os autores optaram por revisar manualmente os artigos utilizando as principais conferências e edições especiais de revistas. Esta etapa foi realizada entre o quarto trimestre de 2009 e o primeiro trimestre de 2010, com refinamentos no último trimestre de 2011. É importante destacar que neste estudo os autores não utilizaram uma

*string* de busca, conforme realizado no estudo de 2010. A revisão resultou em 296 artigos. A Tabela 4 lista os locais de busca utilizados neste estudo e número de artigos encontrados.

Tabela 4 - Lista de locais, acrônimos e número de artigos da RSL de 2012.

Smite et al. 2012		
Local	Acrônimo	Número artigos
<i>International Conference on Software Engineering</i>	ICSE	30
<i>IEEE Software</i>	IEEE Software	17
<i>Communications of the ACM</i>	Comm. of the ACM	6
<i>Information and Software Technology</i>	IST	2
<i>Foundations of Software Engineering</i>	FSE	4
<i>Software Process Improvement and Practice</i>	SPIP	13
<i>International Conference on Global Software Engineering</i>	ICGSE	224
<b>Total geral</b>		<b>296</b>

Assim como ocorreu na primeira revisão sistemática de 2010, mais de 2/3 dos artigos incluídos foram encontrados na ICGSE (Figura 3).

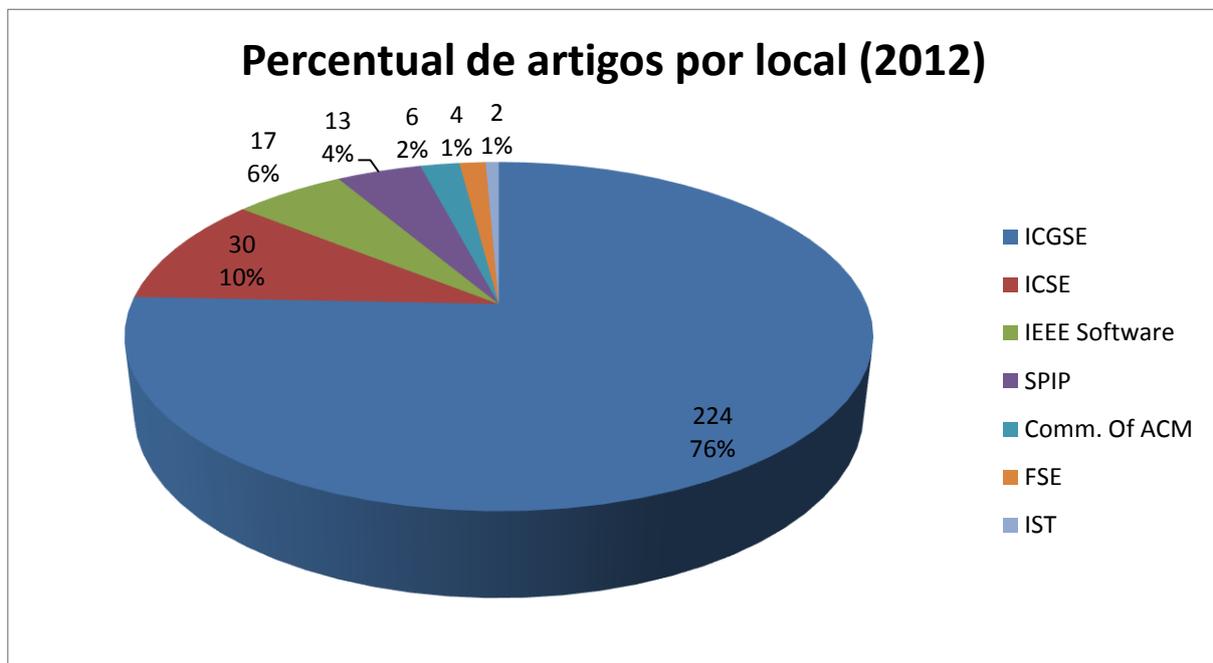


Figura 3 - Percentual de artigos por local.

### 2.3.2 Comparação dos resultados de cada RSL

Após uma análise dos dois estudos, encontramos algumas diferenças em relação ao número de artigos encontrados em cada revisão sistemática. O primeiro estudo selecionou artigos no período de 2000 a 2007 e o segundo estudo selecionou artigos de

2000 a 2011. Para fins de comparação, restringiu-se a análise dos dois artigos para o mesmo período, ou seja, de 2000 a 2007, tornando possível a comparação dos mesmos.

A RSL de 2010 selecionou 59 artigos, enquanto que a revisão de 2012 encontrou 118 artigos no mesmo período. Porém a diferença mais importante foi em relação ao número de artigos selecionados na primeira revisão sistemática que não foram incluídos na segunda revisão sistemática e vice-versa. Conforme o gráfico da Figura 4, no primeiro estudo, 35 artigos não foram incluídos no estudo de 2012. Na RSL de 2012, 94 artigos publicados no período de 2000 a 2007 não foram incluídos na revisão de 2010.

Segundo Smite et al. 2012 [SMI12], os pesquisadores têm enfrentado dificuldades em buscar e encontrar pesquisa relevante, uma vez que as palavras-chave usadas neste campo são muito diversas. Podemos ver claramente na Tabela 5 que existe uma diferença grande quando utilizamos *string* de busca em bases de dados comparada a pesquisa manual.

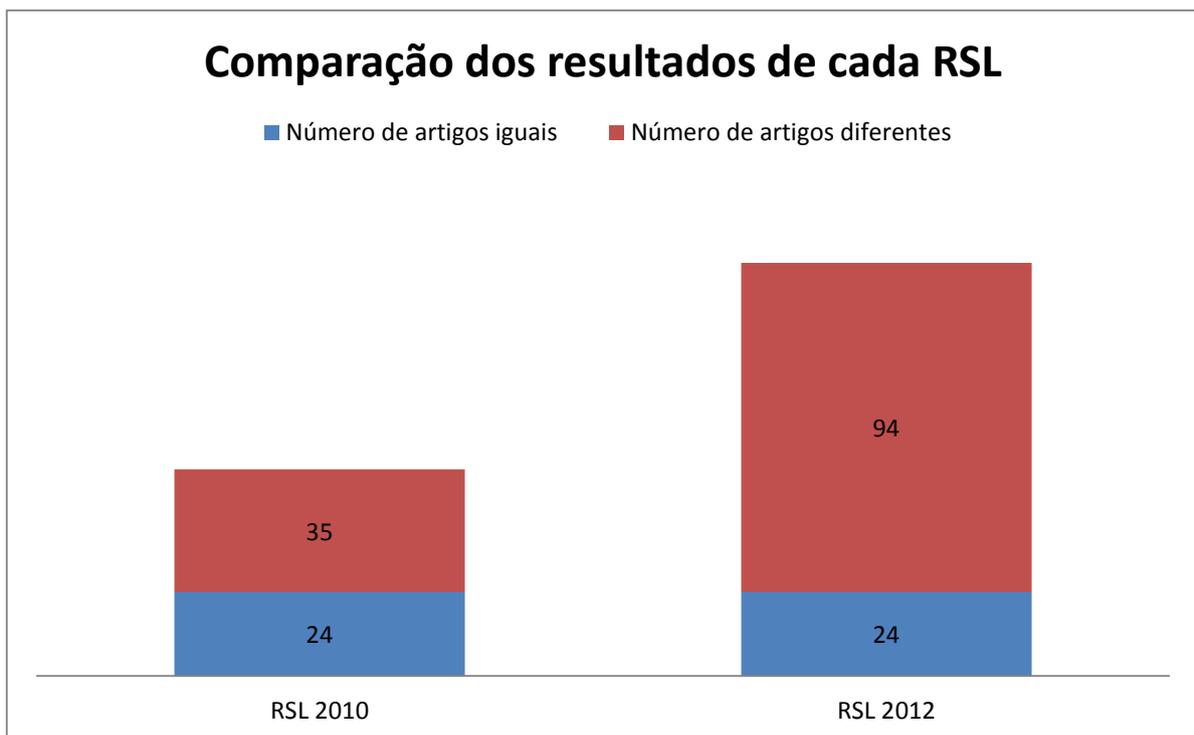


Figura 4 - Comparação dos resultados no período de 2000 a 2007.

A Tabela 5 descreve o resultado da comparação do número de artigos encontrado nos locais de publicação que foram utilizados pelas duas revisões no mesmo período de 2000 a 2007. Evidencia-se, portanto, que a busca manual executada na RSL de 2012 foi mais eficiente, encontrando 118 artigos relevantes, quando comparada a *string* de busca em bases de dados (RSL de 2010) que encontrou 28 artigos nos mesmos locais (conferencias e revistas).

Tabela 5 - Comparação das revisões sistemáticas por local de publicação.

Período de 2000 a 2007			
Local de publicação	Tipo	RSL 2010	RSL 2012
<i>IEEE International Conference on GSE (ICGSE)</i>	Conferência	12	67
<i>International Conference on Software Engineering (ICSE)</i>	Conferência	2	19
<i>Information and Software Technology (IST)</i>	Revista	4	2
<i>Foundations of Software Engineering (FSE)</i>	Conferência	1	1
<i>Software Process Improvement and Practice (SPIP)</i>	Revista	5	6
<i>IEEE Software</i>	Revista	3	17
<i>Communications of the ACM</i>	Revista	1	6
<b>Total</b>		<b>28</b>	<b>118</b>

A revisão sistemática da literatura empírica em DDS ajuda a observar o estado-da-arte, assim como, o estado-da-prática no que se refere ao campo do DDS. Por isso a importância de termos uma base de conhecimento única, com uma terminologia precisa, evitando, portanto, gaps em uma revisão sistemática da área de DDS, tais como os que foram relatados nessa seção.

### 2.3.3 Discussão

Esta seção apresentou uma análise sobre as pesquisas científicas conduzidas na área de sistematização das evidências empíricas em DDS buscando oportunidades de pesquisa e um embasamento teórico para pesquisas futuras. Ao longo do desenvolvimento deste trabalho, foi possível amadurecer o conhecimento sobre como estudos empíricos em DDS podem ser sistematizados de forma clara e intuitiva, trazendo benefícios tanto para os pesquisadores como para os profissionais da indústria que contribuem com o desenvolvimento de pesquisas e relatos práticos sobre os desafios de DDS. Nota-se que ainda há uma escassez de trabalhos abordando esta mesma temática sob a perspectiva adotada nesta pesquisa: sistematização de evidências empíricas em DDS. Além disso, a terminologia aplicada ao DDS não é padronizada e muitas vezes o contexto de um estudo empírico não é descrito de forma clara, gerando, portanto, uma demanda para novos estudos de sistematização das evidências empíricas nesta área.

Com base na análise crítica dos estudos, constatou-se que atualmente o processo de busca por estudos empíricos relacionados com DDS é custoso e muitas vezes não reflete o estado atual da área devido a grande diversidade de termos e conceitos envolvidos. O processo de pesquisa por estudos de caso, soluções, métodos, processos e ferramentas em DDS pode ser melhorado com uso de taxonomias, conforme visto no

estudo de [SMI12]. O método utilizado pelos autores para identificar uma terminologia e para propor uma taxonomia pode ser utilizado em outras áreas da Engenharia de Software. Através de uma investigação da utilização de uma terminologia na literatura é possível identificar semelhanças e diferenças na utilização dos termos de um campo.

Por fim, terminologias com base empírica e taxonomias podem ajudar na maturação de uma área e, conseqüentemente, podem ser utilizadas ou adaptadas em outros domínios ou subáreas de pesquisa em Engenharia de Software.

## **2.4 SISTEMATIZAÇÃO DO CONHECIMENTO EM ENGENHARIA DE SOFTWARE**

Para aprofundar a base teórica desta pesquisa, optou-se por caracterizar o estado da arte referente à sistematização do conhecimento em Engenharia de Software através da realização de um estudo de mapeamento sistemático.

O mapeamento sistemático é um tipo de revisão sistemática, onde se realiza uma revisão mais ampla dos estudos primários, com o propósito de identificar quais evidências estão disponíveis, bem como identificar lacunas no conjunto dos estudos primários e identificar áreas onde mais estudos primários precisam ser conduzidos [KIT07]. O estudo de mapeamento sistemático fornece uma visão geral de uma área de pesquisa, identificando a quantidade, os tipos de pesquisas realizadas, os resultados disponíveis, além das frequências de publicações ao longo do tempo para identificar tendências [PET08]. Segundo Kitchenham e Charters [KIT07], os estudos de mapeamento sistemático em engenharia de software têm sido recomendados principalmente para áreas de pesquisa em que existem poucos estudos primários relevantes e de qualidade. Portanto, o mapeamento sistemático mostrou ser a forma mais adequada aos objetivos do estudo e utilizou-se da mesma metodologia de base da revisão sistemática, com a intenção de ser uma revisão formal, rigorosa e confiável.

Dois pesquisadores participaram do processo de revisão, sendo um aluno de mestrado como principal revisor, o qual foi responsável por desenvolver o protocolo de pesquisa, estratégia de busca, a seleção dos estudos, a avaliação da qualidade dos estudos, a extração e síntese dos dados e os relatórios dos resultados encontrados. O outro revisor foi responsável por validar o protocolo de revisão, monitorar os resultados da busca e o processo de seleção e verificar os dados da avaliação de qualidade e as informações de extração dos dados. A seguir apresentam-se o planejamento e resultados obtidos nesta atividade.

### 2.4.1 Protocolo do mapeamento sistemático

O mapeamento sistemático foi executado seguindo as recomendações documentadas por Kitchenham e Charters [KIT07] e Petersen et al. [PET08]. Desta forma, a condução da revisão seguiu etapas definidas (Figura 5), que serão detalhadas nas próximas seções.

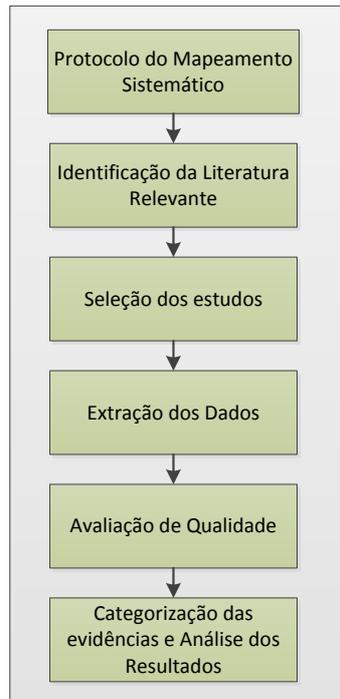


Figura 5 - Etapas do mapeamento sistemático.

Um protocolo de revisão (Apêndice A) que possuía as diretrizes para a condução do mapeamento sistemático foi criado. O objetivo da revisão era responder as seguintes questões de pesquisa:

QP1: Quais são as áreas da Engenharia de Software que apresentam estudos com o objetivo de sistematizar o conhecimento, através de estratégias como taxonomias, esquemas de classificação ou terminologias e quais são os principais estudos publicados na literatura?

QP2: Quais são os desafios e os resultados ou contribuições relatadas nos estudos sobre a utilização de uma ou mais estratégias de sistematização tais como taxonomias, esquemas de classificação ou terminologias?

As estratégias de sistematização descritas na primeira e na terceira questão de pesquisa (QP1 e QP3) foram extraídas dos estudos empíricos analisados na seção 2.2

deste trabalho. O estudo [SMI10] define um esquema de classificação para coleta e sistematização dos dados empíricos gerados por estudos científicos e o estudo [SMI12] apresenta uma taxonomia e uma terminologia unificada para área de DDS. Portanto, para esta revisão, utilizaram-se as três estratégias de sistematização apresentadas nesses estudos.

#### 2.4.2 Identificação da literatura relevante

A estratégia adotada neste estudo foi utilizar uma base de dados que indexa publicações de diversas bases. A base de dados SCOPUS<sup>2</sup> permite pesquisar em diversas bases tais como *IEEE*, *ACM*, *Elsevier* e outras e atualmente possui 49 milhões de publicações [SCO13]. A primeira tarefa na identificação da literatura presente na área é a construção de uma *string* de busca. Desta forma, as estratégias para a identificação dos termos da pesquisa utilizadas foram:

- Derivação de termos utilizados nas questões de pesquisa;
- Listar as palavras-chave dos artigos previamente selecionados e que formam a base deste estudo (SMI10 e SMI12);
- Uso do operador booleano OR para incorporar sinônimos;
- Uso do operador AND para fazer a junção entre as diferentes palavras-chave.

A *string* de busca utilizada para a pesquisa foi:

(("software development" OR "software engineering")  
AND ("taxonomy" OR "classification scheme" OR "terminology"))

O primeiro teste da *string* acima foi realizado na ferramenta SCOPUS e foram retornados 10.532 registros. O motivo para tantos registros serem retornados foi que a *string* de busca está pesquisando no texto completo dos artigos. Portanto, para este estudo, restringiu-se a *string* de busca para pesquisa somente no título, palavras-chaves e resumo. Dessa forma, para esse trabalho é utilizada a seguinte *string* de busca definida para a base de dados SCOPUS:

TITLE-ABS-KEY( ("software development" OR "software engineering")  
AND ("taxonomy" OR "classification scheme" OR "terminology") )

---

<sup>2</sup> <http://www.scopus.com>

A pesquisa na base de dados SCOPUS foi restrita a área “*Computer Science*” para evitar que artigos de outras áreas que não são do interesse dessa revisão fossem retornados. Foram incluídos também restrições de idioma e tipos de documentos que serão descritos na próxima seção.

### 2.4.3 Seleção dos estudos

Para permitir a seleção sistematizada de artigos foram definidos nesta seção alguns critérios de inclusão e exclusão de estudos. O objetivo desses critérios é selecionar apenas estudos que descrevem estratégias de sistematização em Engenharia de Software para uma determinada área, processo, ferramenta, sistema ou tecnologia.

Os artigos devem atender os critérios definidos abaixo para serem incluídos na revisão:

- Artigos publicados em conferência e periódicos;
- Estudos primários ou secundários;
- Artigos na área de Ciência da Computação.
- O idioma dos estudos deve ser somente o inglês.

Os artigos que se enquadrarem em algum dos critérios abaixo serão excluídos da revisão:

- Artigos que não sejam da área de Ciência da Computação;
- Artigos que não sejam em sua totalidade no idioma inglês;
- Artigos cujo foco central não é sistematização de uma determinada área, processo, ferramenta, sistema ou tecnologia;
- Estudos publicados pelo mesmo autor, sobre o mesmo tópico de pesquisa (somente estudo mais completo será selecionado e os demais serão excluídos);
- Estudos que expressam a opinião pessoal de alguém e que não dependem de trabalhos relacionados e metodologias de pesquisa;
- Estudos que apenas sugerem estratégias de sistematização, tais como, taxonomia, esquema de classificação e terminologia, porém não utilizam as mesmas no estudo em questão;
- Estudos que mencionam ou utilizam estratégias de sistematização existentes na literatura, porém não desenvolvem nenhum trabalho de melhoria ou adaptação das mesmas.

- Tutoriais, *short papers* e palestras.

O processo de seleção de estudos consistiu em duas fases. Na primeira fase, para cada artigo pesquisado, foi analisado o título, o resumo e as palavras chaves. Os artigos relevantes foram armazenados para uma avaliação posterior. Na segunda fase cada estudo relevante foi lido na íntegra para determinar se era um estudo primário desta revisão.

#### 2.4.4 Avaliação da qualidade

A qualidade e importância dos artigos com relação à questão de pesquisa serão classificadas conforme os itens abaixo:

- O estudo descreve problemas ou desafios relacionados à sistematização do conhecimento?
- O estudo descreve detalhadamente ao menos uma das estratégias de sistematização para contornar os problemas ou desafios relatados?
- O estudo apresenta evidências empíricas para avaliação da estratégia de sistematização proposta?

#### 2.4.5 Extração dos dados

Para a extração dos dados foi utilizado uma ficha de leitura desenvolvida no MS Excel<sup>3</sup>. Os itens da ficha foram selecionados conforme alinhamento das questões de pesquisa. A ficha possuía os seguintes itens:

- Identificação (ID)
- Título
- Autor
- Ano
- Veículo (onde foi publicado)
- Volume
- Edição
- Início e fim da página
- Estratégia (taxonomia, esquema de classificação, terminologia)
- Área de conhecimento do SWEBOK

---

<sup>3</sup> <http://office.microsoft.com/pt-br/excel/>

- Descrição do problema relacionado com sistematização
- Objetivo do estudo
- Contribuição, Evidências
- Status (Incluso ou Não Incluso)
- Justificativa (referente ao status)

#### 2.4.6 Execução

Nessa seção apresentamos o resultado da execução do protocolo do mapeamento sistemático e número de artigos resultante de cada fase. A partir do protocolo de pesquisa definido, o mapeamento sistemático foi executado na base de dados SCOPUS. Os dados resultantes da pesquisa foram exportados para arquivos texto no formato CSV. A primeira execução da *string* de busca retornou 539 artigos.

Na fase inicial todos os artigos retornados pela *string* de busca tiveram seu título e resumo analisados considerando os critérios de seleção e exclusão definidos anteriormente. Após essa análise o número de artigos selecionados foi reduzido para 94.

Na segunda fase todos os artigos selecionados na fase anterior foram lidos integralmente e um conjunto de informações foi extraído de cada estudo, conforme descrito na seção 2.4.5. Os artigos que não se enquadraram nos critérios de seleção foram removidos da seleção. Após essa análise o número de artigos selecionados foi reduzido para 43. A Tabela 6 apresenta o número de artigos resultantes em cada etapa executada.

Tabela 6 - Número de artigos por etapa.

Fases	Número de artigos
<b>Execução da <i>string</i> de busca</b>	539
<b>Fase 1</b>	94
<b>Fase 2</b>	43

#### 2.4.7 Análise dos resultados

Esta seção apresenta uma análise quantitativa e qualitativa dos estudos identificados no mapeamento sistemático desenvolvido, assim como identifica os principais problemas sobre sistematização relatados na literatura e as estratégias utilizadas para contornar tais problemas.

A lista com todos os artigos analisados encontra-se no Apêndice B deste trabalho. Para referenciá-los utilizou-se um padrão de número específico com o objetivo de diferenciar das outras referências deste trabalho.

#### 2.4.7.1 Análise quantitativa

Nesta seção apresentamos uma análise quantitativa sobre as características dos estudos, tais como, publicações por ano, publicações por área da Engenharia de Software e publicações por estratégia de sistematização.

##### 2.4.7.1.1 Distribuição dos estudos por ano

A distribuição dos artigos ao longo dos anos é apresentada na Figura 6. No ano de 2013 os artigos foram analisados até o mês de Junho (mês que a revisão foi executada).

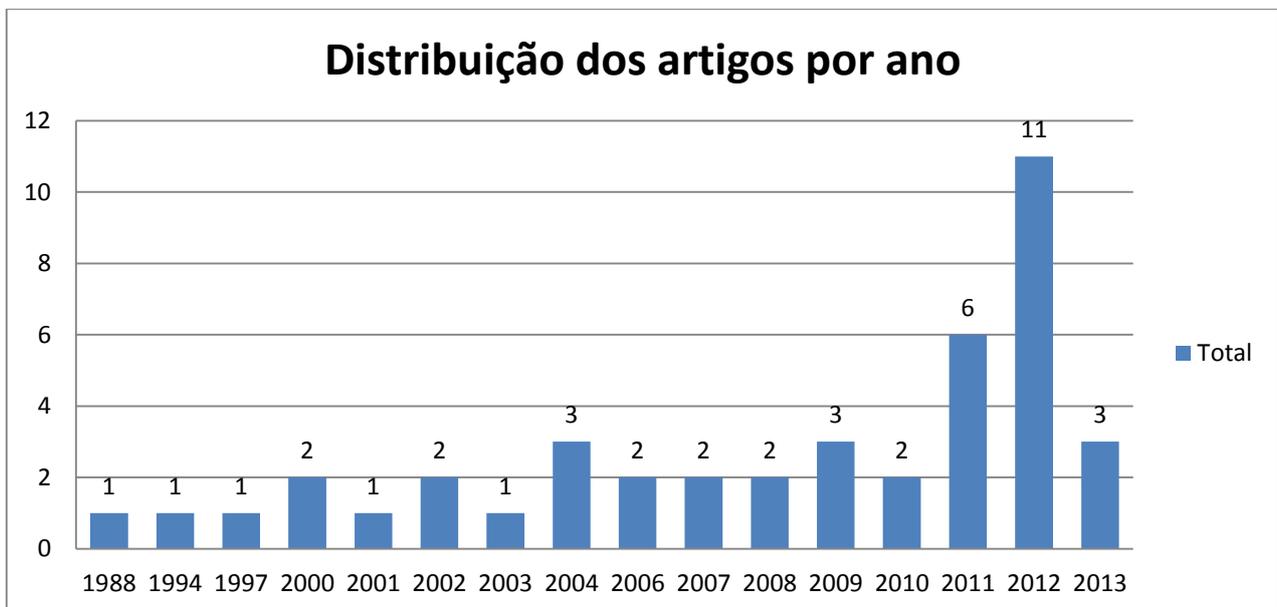


Figura 6 - Ano de publicação dos artigos identificados.

##### 2.4.7.1.2 Distribuição dos estudos por década

A distribuição dos estudos por década é apresentada na Figura 7. Destacam-se as décadas de 2000 e 2010, totalizando 40 dos 43 estudos selecionados.

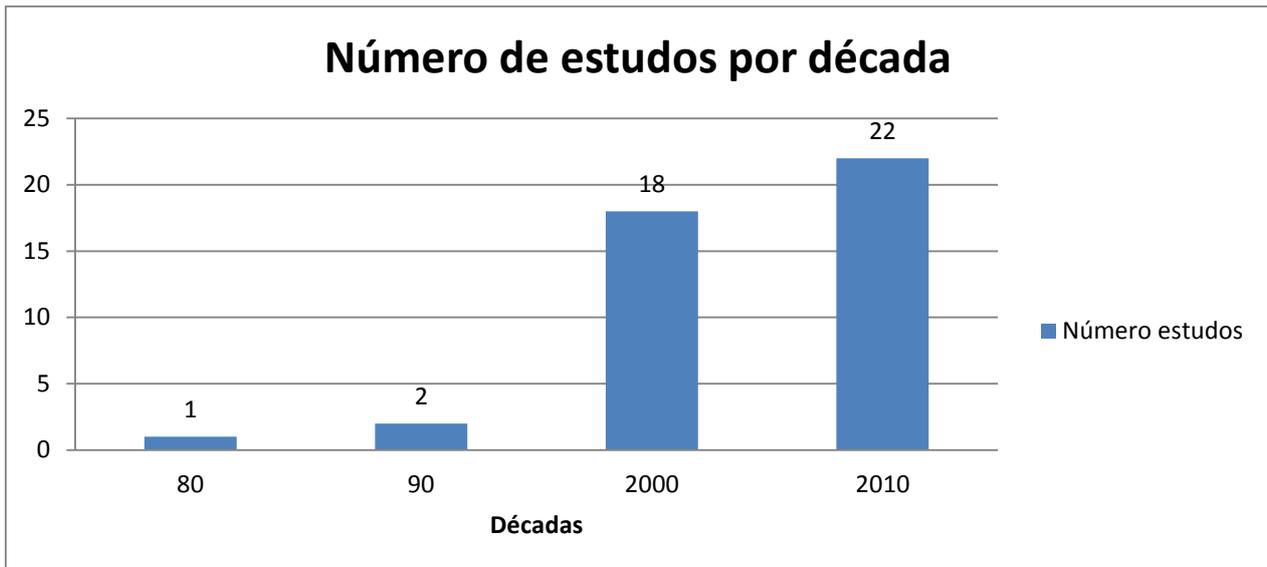


Figura 7 - Número de estudos por década.

#### 2.4.7.1.3 Distribuição por área de conhecimento do SWEBOK

O guia *Software Engineering Body of Knowledge*, ou SWEBOK [ABR04], produzido pela *IEEE Computer Society*, é um guia para o corpo de conhecimento da Engenharia de Software. Este documento foi revisado por profissionais de todo o mundo, inclusive o Brasil, e mapeia o conhecimento em Engenharia de Software. O guia foi criado com os seguintes objetivos:

- Promover uma visão consistente da Engenharia de Software para o mundo;
- Esclarecer e definir o limite, da Engenharia de Software em relação a outras disciplinas, como Ciência da Computação, Gerenciamento de Projetos, Engenharia da Computação e Matemática;
- Caracterizar os conteúdos da disciplina de Engenharia de Software;
- Permitir um acesso por tópicos para o corpo de conhecimento de Engenharia de Software;
- Fornecer uma base para o desenvolvimento curricular, para a certificação individual e material de licenciamento.

O guia SWEBOK é dividido em dez áreas de conhecimento e cada área é subdividida em subáreas. As dez áreas de conhecimento são:

- Requisitos de Software
- Design de Software
- Construção de Software
- Teste de Software

- Manutenção de software
- Gerência de Configuração de Software
- Gerência de Engenharia de Software
- Processos de Engenharia de Software
- Ferramentas e Métodos de Engenharia de Software
- Qualidade de Software

Todos os artigos selecionados foram classificados em umas das dez áreas do corpo de conhecimento do SWEBOK. A distribuição dos artigos por área é apresentada na Figura 8. Não foram encontrados estudos nas áreas de Teste de Software, Manutenção de software, Gerência de Configuração de Software e Gerência de Engenharia de Software.

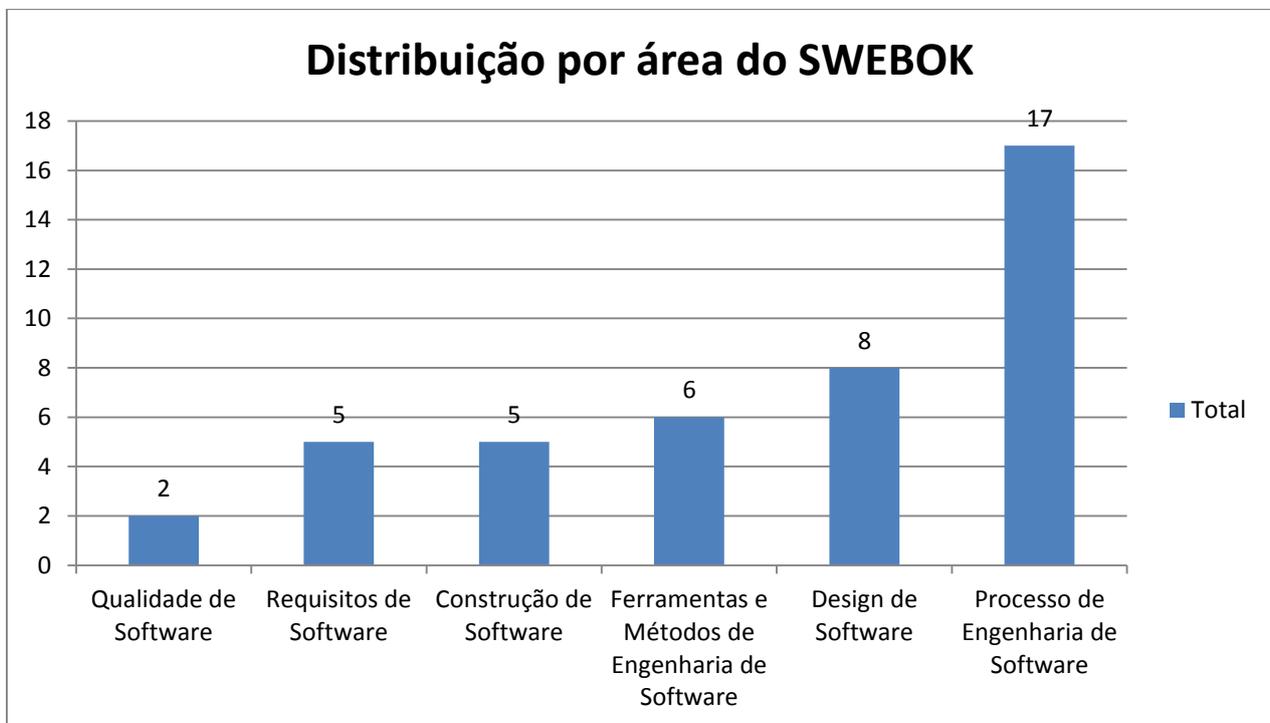


Figura 8 - Distribuição dos artigos por área do SWEBOK.

Conforme podemos ver na Figura 9, a área de Processo de Engenharia de Software destaca-se como a área de conhecimento do SWEBOK com maior demanda de sistematização do conhecimento, com 39% dos artigos. A área de Design de Software ficou em segundo lugar, com 18% dos artigos e a área de Ferramentas e Métodos de ES ficou com 14%. As áreas de Construção de Software e Requisitos de Software também apresentaram um número considerável de artigos, chegando a 12% cada. Por último a área de Qualidade de Software apresentou apenas dois estudos, totalizando 5% dos artigos.

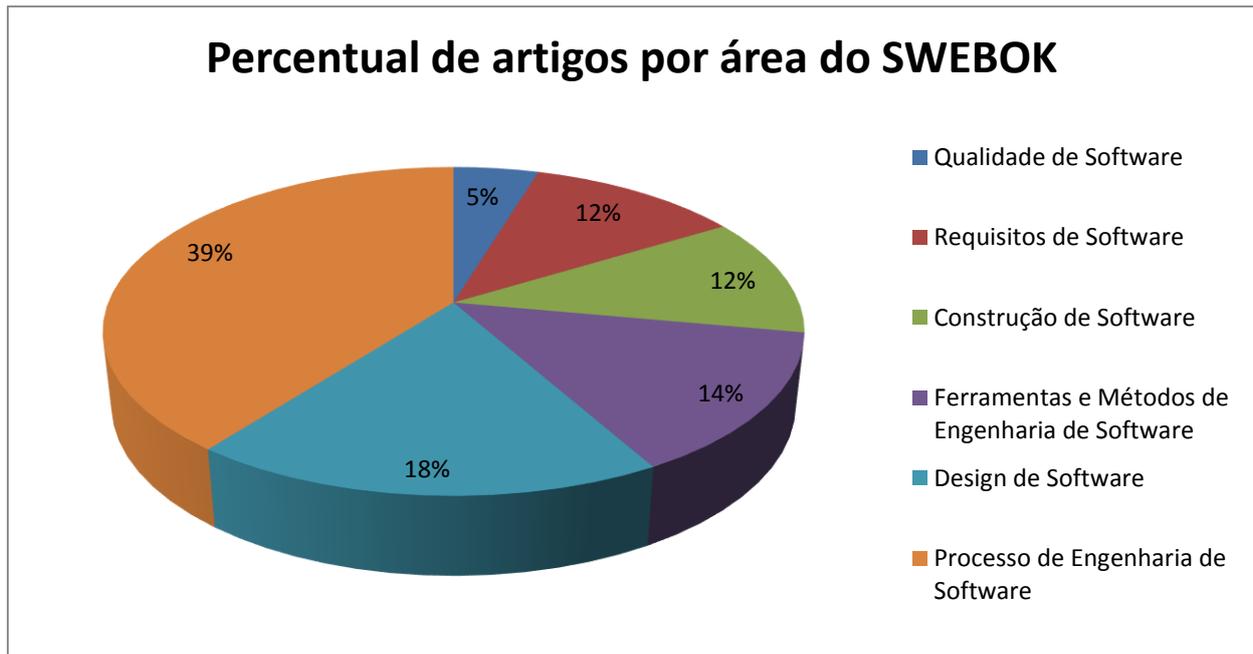


Figura 9 - Percentual de artigos por área do SWEBOK.

#### 2.4.7.1.4 Distribuição por estratégia de sistematização

As estratégias de sistematização definidas na *string* de busca foram utilizadas para classificar os artigos selecionados. Durante a análise detalhada os artigos foram classificados em três categorias: taxonomia, esquema de classificação e terminologia. Alguns artigos apresentaram uma taxonomia e uma terminologia e, portanto, foram agrupados em uma quarta categoria conforme podemos ver na Figura 10.

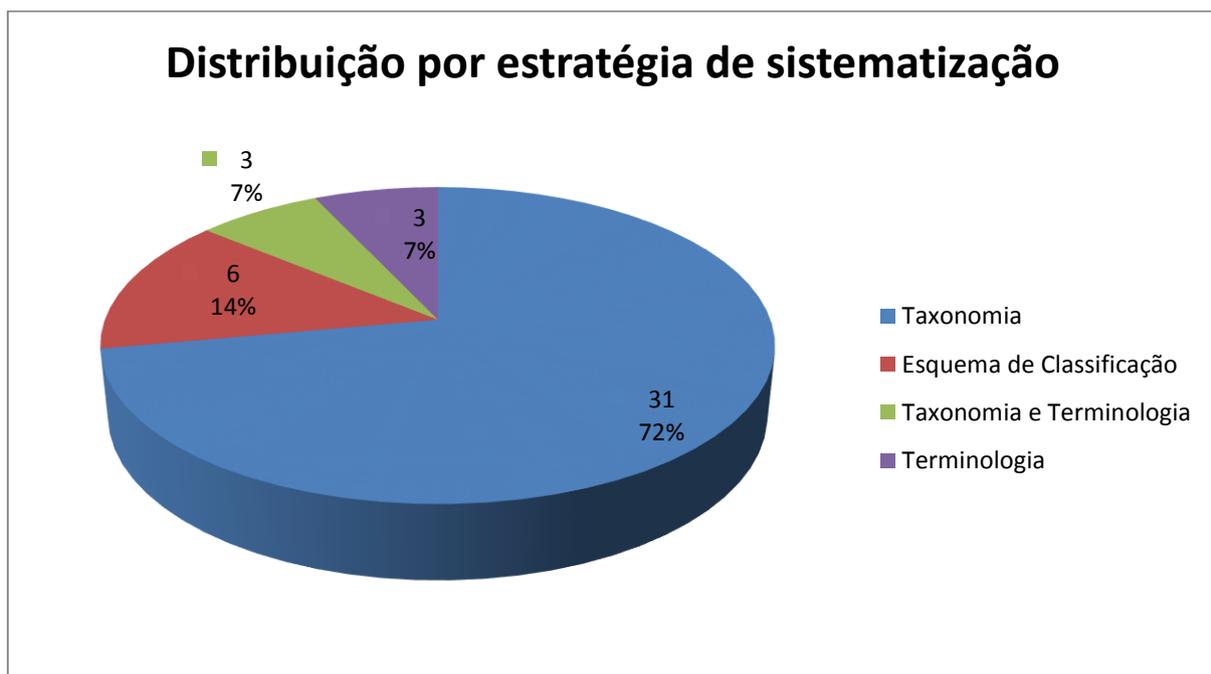


Figura 10 - Distribuição por estratégia de sistematização.

Observamos que mais de 70% dos artigos utilizaram taxonomias como estratégia de sistematização primária. A segunda estratégia mais utilizada foi o esquema de classificação, com 14% dos estudos. Somente três artigos apresentaram duas estratégias no mesmo estudo, ou seja, uma taxonomia e uma terminologia, como resposta aos problemas de padronização dos termos de uma determinada área da ES. Por fim, três estudos apresentaram apenas uma terminologia para sistematização de uma área de conhecimento da ES.

#### 2.4.7.1.5 Distribuição ao longo dos anos por estratégia de sistematização

A Figura 11 mostra a distribuição dos artigos ao longo dos anos por estratégia de sistematização. Taxonomia é a estratégia mais utilizada ao longo dos anos, com um forte crescimento do número de estudos publicados no ano de 2012.

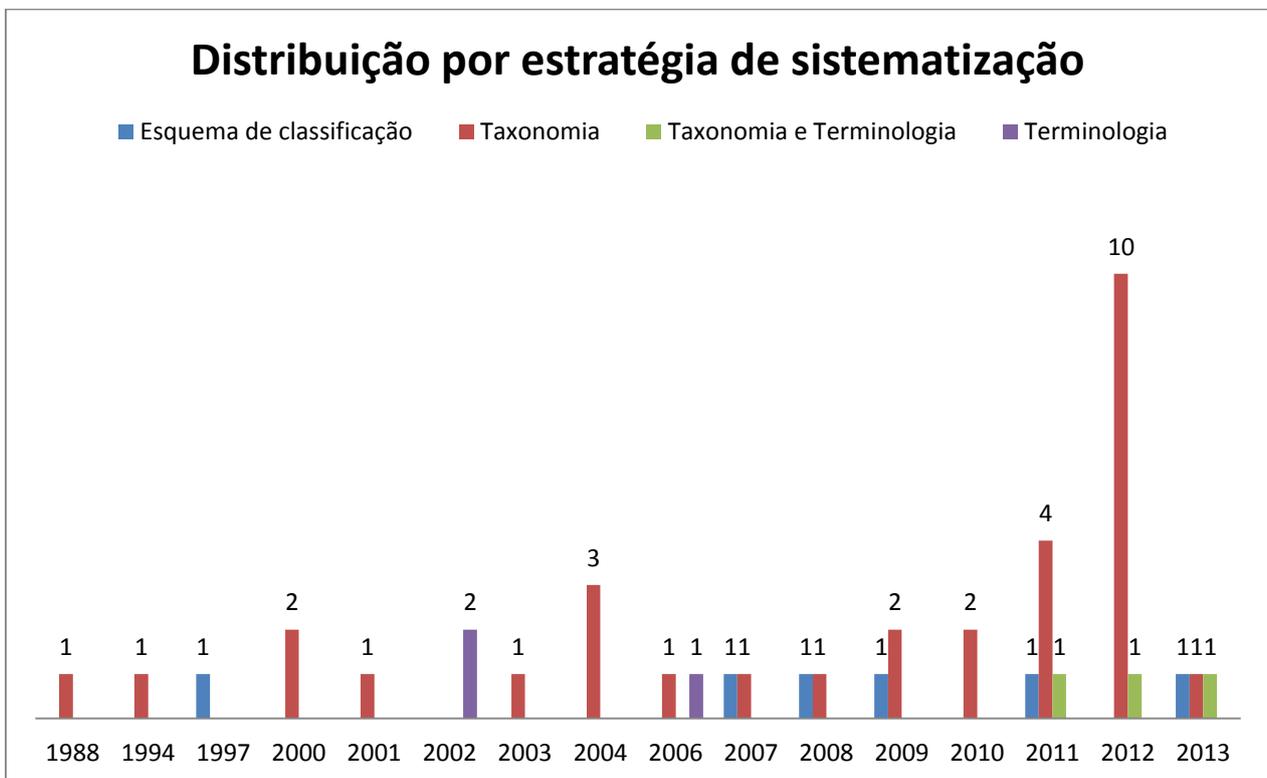


Figura 11 - Distribuição por estratégia ao longo dos anos.

#### 2.4.7.2 Análise qualitativa

Esta seção apresenta uma análise qualitativa e tem por objetivo caracterizar os estudos selecionados, com o propósito de identificar o conteúdo de cada artigo em vista a responder as questões de pesquisa definidas no protocolo do mapeamento sistemático.

### Questão de pesquisa QP1:

O mapeamento sistemático retornou estudos em diversas áreas da ES, conforme mostramos na Figura 8, respondendo, portanto, a primeira questão de pesquisa (QP1). A distribuição dos artigos por área de conhecimento do SWEBOK e por estratégia de sistematização utilizada nos estudos encontra-se na Tabela 7.

Tabela 7 - Distribuição dos artigos por área e estratégia de sistematização.

Área do SWEBOK	Taxonomia	Esquema de Classificação	Terminologia	Taxonomia & Terminologia	Total
<b>Processo de Engenharia de Software</b>	[CHE12], [FOR08], [GFP12], [GUM06], [KAJ01], [MAL11], [PRE94], [RAH12], [ROB12], [SAV07], [STR12], [WAL13]	[PET11], [SMI08]	[GRI06]	[SMI12], [VAS11]	17
<b>Design de Software</b>	[HOF11], [MEH00], [ROM11], [TAI11], [WUY09]	[OLU07], [SAK09]	[CRN02]		8
<b>Ferramentas e Métodos de Engenharia de Software</b>	[ALI10], [DAV88], [DEL04], [NOV10], [SUL04], [UTT12]				6
<b>Construção de Software</b>	[CHA12], [MON12], [YAN09]	[MEH13]		[FRE13]	5
<b>Requisitos de Software</b>	[GHA12], [NUR04], [RAM12]	[ZAV97]	[KHW02]		5
<b>Qualidade de Software</b>	[LAI00], [MAN03]				2
<b>Total</b>					<b>43</b>

### Questão de pesquisa QP2:

Nas próximas seções descrevemos os desafios e contribuições relatados nos estudos relacionados com as áreas da ES classificados de acordo com o Corpo de

Conhecimento de Engenharia de Software (SWEBOK), respondendo, portanto, a questão de pesquisa Q2.

#### 2.4.7.2.1 Processo de Engenharia de Software

Segundo SEWBOK [SEW04], a área de conhecimento em Processo de Engenharia de Software pode ser analisada sob duas perspectivas. A primeira engloba as atividades técnicas e gerenciais, dentro do processo do ciclo de vida de software, que são executadas durante aquisição, desenvolvimento, manutenção e aposentadoria do software. A segunda está preocupada com a definição, execução, avaliação, medição, gestão, mudança e a melhoria do processo de ciclo de vida do software. A Tabela 8 resume os desafios e as contribuições de cada estudo desta área.

Tabela 8 - Sistematização em Processo de Engenharia de Software.

ID	Contexto	Estratégia	Desafio	Resultado ou contribuição
[CHE12]	<i>Knowledge Management in the Software Industry Sector</i>	Taxonomia	Compreender como o conhecimento pode ser modelado e classificado de forma que ele possa ser eficientemente compartilhado na indústria.	Apresenta uma taxonomia que define os tipos de conhecimento e suas representações (tácito ou explícito) que pode ser utilizada tanto no meio acadêmico como pela indústria.
[FOR08]	<i>Software Types</i>	Taxonomia	Melhorar a pesquisa empírica em software de uma forma mais sistemática para identificar os tipos de software para o qual se aplica a evidência empírica.	A taxonomia proposta pode ser usada para categorizar estudos empíricos e outras evidências obtidas por pesquisadores, de modo que a aplicabilidade da evidência possa ser entendida de forma mais sistemática.
[GFP12]	<i>Software Platform Strategy</i>	Taxonomia	Compreender os problemas e os desafios que as organizações enfrentam quando adotam estratégias de plataforma de software.	Este estudo fornece uma taxonomia abrangente dos desafios enfrentados quando uma organização decide adotar plataformas de software. O estudo identifica classes de desafios que

ID	Contexto	Estratégia	Desafio	Resultado ou contribuição
				podem ajudar os pesquisadores e profissionais a identificar soluções práticas e ferramentas de apoio.
<b>[GRI06]</b>	<i>Software Effort Estimation</i>	Terminologia	A falta de uma terminologia precisa para estimativa de esforço de software é um obstáculo importante para a melhoria da precisão de estimativas. O termo "estimativa de esforço" é frequentemente usado sem o esclarecimento do seu significado e que a precisão da estimativa é frequentemente avaliada sem garantir que a estimativa e o esforço real são comparáveis.	O estudo propõe dois <i>guidelines</i> para melhorar a terminologia e o processo de estimativa de software. Os <i>guidelines</i> são destinados a todos que utilizam terminologias para estimativa de custo de software, incluindo autores, profissionais, pesquisadores e revisores.
<b>[GUM06]</b>	<i>Distribution Dimensions in Global Software Engineering</i>	Taxonomia	Compreender o fenômeno de distribuição dos projetos de desenvolvimento de software global.	Taxonomia preliminar que identifica as dimensões de distribuição, que podem ser utilizados pelos desenvolvedores de software para visualizar as dimensões de distribuição do seu projeto.  A taxonomia ajuda a compreender a distribuição física da equipe de projeto.
<b>[KAJ01]</b>	<i>Problem Management Process</i>	Taxonomia	Processo de gerenciamento de problemas é um dos muitos processos realizados dentro da manutenção corretiva de	O estudo propõe uma taxonomia para classificar as atividades essenciais dentro de um processo de gerenciamento de problemas.

ID	Contexto	Estratégia	Desafio	Resultado ou contribuição
			software, porém existem muitas categorias de manutenção, tornando o processo de manutenção impreciso e subjetivo.	
[MAL11]	<i>Agent based System Complexity Measurement</i>	Taxonomia	Falta de consenso sobre o termo complexidade em projetos e não existe um entendimento comum de complexidade e seus indicadores em sistemas baseados em agentes.	Apresenta uma taxonomia para medir a complexidade dos sistemas baseados em agentes, com base em uma mistura de componentes de gerenciamento de projetos e componentes de engenharia de software.
[PET11]	<i>Software Productivity Measurement</i>	Esquema de Classificação	Identificar e agregar os elementos de medição da produtividade do desenvolvimento de software.	O estudo apresenta um esquema de classificação para a pesquisa empírica sobre a produtividade de software. O mapeamento permite identificar às diferentes abordagens de produtividade e ajuda na sistematização dos estudos.
[PRE94]	<i>Software Engineering and Human-Computer Interaction Measurement</i>	Taxonomia	Compreender as diferenças e semelhanças das abordagens e técnicas de medição utilizadas nas áreas de Engenharia de Software e Interação Humano-Computador.	Uma taxonomia para descrever as abordagens e técnicas de medição utilizadas nas duas áreas. A taxonomia pode ser usada para explicar o processo de medição, examinar pesquisas anteriores em ambos os campos e facilitar a comunicação entre as duas disciplinas, fornecendo um vocabulário comum.
[RAH12]	<i>Software Process</i>	Taxonomia	Frameworks diferem em vários aspectos e,	Criação de uma taxonomia para comparar os frameworks

ID	Contexto	Estratégia	Desafio	Resultado ou contribuição
	<i>Improvement Frameworks</i>		portanto, são difíceis de serem comparados, interpretados e compreendidos.	de melhoria de processo em engenharia de software. A Taxonomia pode ser usada para apoiar a compreensão e comparação de diversos frameworks de qualidade de software.
<b>[ROB12]</b>	<i>Software Team Process</i>	Taxonomia	O conceito de processo de equipe nunca foi definido claramente na literatura de Engenharia de Software.	O estudo propõe uma taxonomia para ajudar futuros pesquisadores a reportar o resultado de suas pesquisas de forma mais clara. A taxonomia fornece um vocabulário comum que também pode auxiliar no entendimento dos resultados de outros estudos sobre o mesmo tema.
<b>[SAV07]</b>	<i>Security Metrics</i>	Taxonomia	<p>O campo de definição de métricas de segurança ainda é jovem.</p> <p>A prática atual da segurança da informação é um campo altamente diversificado e não existem abordagens holísticas e amplamente aceitas nesta área.</p> <p>Para obter evidência da segurança de diferentes produtos ou organizações, abordagens sistemáticas para a segurança de medição são necessárias.</p>	O estudo apresenta uma taxonomia para apoiar o desenvolvimento de métricas de segurança viáveis, juntamente com um levantamento das métricas de segurança emergentes a partir das perspectivas acadêmicas, governamentais e industriais.
<b>[SMI08]</b>	<i>Empirical</i>	Esquema de	Os cenários em GSE são	O estudo apresenta um

ID	Contexto	Estratégia	Desafio	Resultado ou contribuição
	<i>Research in Global Software Engineering</i>	Classificação	diversas, o que funciona em um contexto pode não se aplicar diretamente em outro e também é necessário compreender como resultados empíricos devem ser relatados para serem úteis para ambos profissionais e pesquisadores da área.	esquema de classificação que pode ser utilizado para compreender o contexto dos estudos existentes e orientar o processo de informação e estruturação de estudos futuros.
<b>[SMI12]</b>	<i>Sourcing in Global Software Engineering</i>	Taxonomia & Terminologia	A terminologia em projetos globais é muito diversificada (muitos sinônimos usados para descrever os mesmos fenômenos), muitas vezes confusa (os mesmos termos usados para descrever fenômenos diferentes) e ambígua (poucos termos usados para descrever vários fenômenos).	Os autores propõem uma terminologia unificada e uma taxonomia para GSE para categorizar os termos com base em relacionamentos de generalização-especialização e também para classificar e mapear o conhecimento existente da área. A contribuição visa futuros pesquisadores, que irão publicar ou sintetizar novos trabalhos e profissionais que estão interessados em estudos com base empírica em GSE.
<b>[STR12]</b>	<i>Dependency analysis in Agile Software Engineering</i>	Taxonomia	Caracterizar que tipo de dependência ocorre em projetos de desenvolvimento de software ágeis.	Este trabalho contribui para academia, proporcionando uma taxonomia de tipos de dependências que ocorrem na área de desenvolvimento ágil de software.
<b>[VAS11]</b>	<i>Uncertainty in Software Development</i>	Taxonomia & Terminologia	No contexto de desenvolvimento de software, o termo incerteza e os demais termos associados a ele,	Apresenta uma terminologia e uma taxonomia para classificar as fontes de incerteza durante o processo de desenvolvimento e

ID	Contexto	Estratégia	Desafio	Resultado ou contribuição
			tais como, erro, risco e a ignorância, são definidos e interpretados de forma diferente por diversos autores.	modelagem de sistemas. A compreensão das fontes de incerteza pode ajudar o desenvolvedor a projetar o custo do software e a estimativa de programação de forma mais precisa.
[WAL13]	<i>Error Abstraction Process</i>	Taxonomia	Identificar falhas e erros no início do ciclo de vida de desenvolvimento de software.	O estudo propõe uma taxonomia com classes de erros de requisitos para que os desenvolvedores possam reconhecer as categorias de problemas e identificar falhas quando estão revisando um artefato de software.

#### 2.4.7.2.2 Design de Software

O guia SWEBOK [ABR04] define a área de conhecimento em Design de Software como o processo de definição da arquitetura, componentes, interfaces e outras características de um sistema ou componente. O guia descreve que o design de software desempenha um papel importante no desenvolvimento de software, pois permite que os engenheiros de software produzam modelos que formam “*blueprints*” da solução a ser desenvolvida.

Segundo [OLU07], uma abordagem bem aceita para compartilhar experiências entre a comunidade de desenvolvedores de software é através da utilização de padrões de software, porém não existe uma organização unificada dos padrões orientados a agentes, e como resultado, o reuso de padrões orientados a agentes é limitado. A indústria de software está atualmente experimentando uma mudança de paradigma para a computação em nuvem e software baseado em web [TAI11]. No entanto, com a quantidade de serviços de computação de nuvem crescendo rapidamente, a necessidade de uma taxonomia aumenta [HOF11]. A Tabela 9 resume os desafios e as contribuições dos estudos encontrados desta área.

Tabela 9 - Sistematização em Design de Software.

ID	Contexto	Estratégia	Desafio	Resultado ou contribuição
[CRN02]	<i>Component-based Software Engineering</i>	Terminologia	ES baseada em componentes utiliza conceitos que ainda não estão totalmente formalizados e termos que não são claramente distinguidos. Por exemplo, o termo "componente" e "interface" não foram formalmente especificados.	O estudo apresenta uma terminologia comum, esclarece os termos utilizados na área de Engenharia de Software baseada em Componentes e discute as relações entre eles.
[HOF11]	<i>Cloud Computing Services</i>	Taxonomia	Comparar os serviços de computação em nuvem.	Este artigo analisa os serviços de computação em nuvem disponíveis e identifica e organiza as suas principais características através de uma taxonomia.
[MEH00]	<i>Software Connectors</i>	Taxonomia	Não existe um entendimento claro sobre os conectores de software. O resultado é a utilização inconsistente destes conectores, devido à falta de compreensão de como eles podem ser compostos em interações mais complexas.	Apresenta um framework de classificação e uma taxonomia de conectores de software. A taxonomia pode ser usada tanto para compreender os conectores de software existentes, como para sugerir novos conectores.
[OLU07]	<i>Agent-oriented patterns</i>	Esquema de Classificação	Encontrar padrões orientados a agentes de uma forma	O estudo descreve um esquema de classificação que fornece uma plataforma

ID	Contexto	Estratégia	Desafio	Resultado ou contribuição
			sistemática e organizada.	para analisar, classificar e descrever padrões orientados a agentes.
[ROM11]	<i>Rich-User-Interface (RUI) components</i>	Taxonomia	Não existe um critério de classificação homogêneo para componentes RUI, o que dificulta a seleção destes componentes.	Apresenta uma taxonomia baseada no estudo dos recursos relevantes de domínio da interface do usuário e na opinião dos especialistas no assunto. Esta taxonomia permite categorizar elementos estruturais e de controle que fazem parte das interfaces de usuário Web.
[SAK09]	<i>Autonomous software agent architectures</i>	Esquema de Classificação	Classificar os diferentes tipos de agentes abaixo da mesma raiz.	Este artigo descreve uma proposta de classificação para arquitetura de agentes autônomos. Ele baseia-se em várias classificações existentes, mas fornece uma classificação abrangente, que leva em conta os diversos aspectos da pesquisa em agentes.
[TAI11]	<i>Web-based Computing Systems</i>	Taxonomia	O cenário da computação em nuvem é muito diversificado e é difícil compreender as dimensões e características dos sistemas de computação baseados na web.	Propõe uma taxonomia para classificar os sistemas de computação baseados na web e criar uma nova estrutura para facilitar a discussão sobre computação em nuvem.
[WUY09]	<i>Software Privacy Solutions</i>	Taxonomia	Muitas soluções de privacidade estão surgindo, porém não	Este artigo propõe uma definição operacional para privacidade e apresenta uma

ID	Contexto	Estratégia	Desafio	Resultado ou contribuição
			existe uma classificação clara das soluções de privacidade de software.	taxonomia de privacidade, que classifica os objetivos de privacidade.

#### 2.4.7.2.3 Construção de Software

A área de conhecimento em Construção de Software refere-se à criação detalhada de software através da combinação de codificação, verificação, testes unitários, testes de integração e depuração. Esta área de conhecimento está ligada a todos as outras áreas do SWEBOOK, porém, mais fortemente com as áreas de Design de Software e de Teste de Software. Isso ocorre porque o processo de construção de software envolve diretamente as atividades de design e teste de software [ABR04].

Com a ampla disseminação da programação orientada a aspectos, mais aspectos são descobertos, projetados e desenvolvidos [YAN09]. O Desenvolvimento de Software Orientado a Aspectos é uma abordagem de Engenharia de Software que permite a identificação, separação e encapsulamento das preocupações (*concerns*) que atravessam a modularização primária de um sistema [MEHM13]. Segundo [YAN09], devido à diversidade de domínios de aplicação, os mecanismos concretos para especificação e implementação de aspectos variam muito, o que torna difícil compará-los. A Tabela 10 resume os desafios e as contribuições de cada estudo desta área.

Tabela 10 - Sistematização em Construção de Software.

ID	Contexto	Estratégia	Desafio	Resultado ou contribuição
[CHA12]	<i>Policy Modeling in Software Engineering</i>	Taxonomia	Compreender o conceito de diretriz ( <i>policy</i> ) em sistemas de software corporativos e identificar as principais áreas de análise de diretrizes.	O estudo apresenta uma taxonomia que classifica as diretivas relacionadas com sistemas corporativos de TI em quatro áreas principais: segurança, processos de negócios, políticas regulatórias e políticas de design.
[FRE13]	<i>Self-healing and self-</i>	Taxonomia & Terminologia	O número de estudos em tecnologias <i>self-healing e</i>	Este estudo apresenta uma terminologia e uma taxonomia

ID	Contexto	Estratégia	Desafio	Resultado ou contribuição
	<i>repairing technologies</i>		<i>self-repairing</i> tem aumentado muito, porém pesquisadores às vezes usam o mesmo termo com diferentes significados ou vários termos para referir-se ao mesmo significado.	para esclarecer os termos utilizados neste campo de pesquisa. O objetivo é fornecer uma terminologia unificada e uma taxonomia que sirva como base para estudos futuros.
[MEH13]	<i>Model-driven Code Generation</i>	Esquema de Classificação	Existe a necessidade de uma investigação mais sistemática sobre o tema.	O estudo apresentou um esquema de classificação, onde classificaram os artigos em três categorias: (1) área de foco, (2) tipo de contribuição e (3) tipo de pesquisa.
[MON12]	<i>Software Development API</i>	Taxonomia	Determinar que tipos de diretrizes de APIs existem e como podemos classificar essas diretrizes. Ou seja, determinar a natureza das diretrizes na documentação de APIs.	Uma taxonomia abrangente de diretrizes que estão presentes na documentação de APIs para melhorar a compreensão da documentação de APIs.
[YAN09]	<i>Classification of aspects</i>	Taxonomia	A especificação e implementação de aspectos variam muito, o que torna difícil compará-los.	O estudo apresenta uma taxonomia para classificação de diferentes aspectos. A taxonomia caracteriza aspectos de várias dimensões no nível mais abstrato.

#### 2.4.7.2.4 Requisitos de Software

Segundo o guia SWEBOK [ABR04], a área de conhecimento Requisitos de Software está preocupada com a licitação, análise, especificação e validação de requisitos de software. É amplamente reconhecido dentro da indústria de software que os projetos de ES são extremamente vulneráveis quando estas atividades são mal realizadas. O guia SWEBOK descreve que um requisito é definido como uma propriedade

que deve ser exibida a fim de resolver algum problema no mundo real. A área de Requisitos de Software apresentou estudos com as três estratégias de sistematização (taxonomias, esquemas de classificação e terminologias). Os estudos apresentaram desafios com a terminologia e também com a compreensão e caracterização da área. A Tabela 11 resume os desafios e as contribuições de cada estudo desta área.

Tabela 11 - Sistematização em Requisitos de Software.

ID	Contexto	Estratégia	Desafio	Resultado ou contribuição
[GHA12]	<i>Enterprise Systems Requirements</i>	Taxonomia	Compreender o espaço de requisitos funcionais no domínio de <i>Enterprise Systems</i> .	O estudo fornece uma caracterização precisa na forma de uma taxonomia de requisitos funcionais e mostra como os requisitos funcionais podem ser distribuídos entre as categorias da taxonomia.
[KHW02]	<i>Software Specifications Evaluation Criteria</i>	Terminologia	Existe uma falta de consistência na definição de propriedades dos diferentes critérios de especificações e critérios de avaliação de técnicas de especificação. A implicação desta falta de diferenciação entre os dois conceitos é que os critérios utilizados em um estudo podem não ser aplicáveis para o outro.	Este artigo apresenta uma terminologia unificada de critérios de especificação e também mapeia as propriedades de critérios de vários pesquisadores na terminologia proposta.
[NUR04]	<i>Requirements volatility in Software Development</i>	Taxonomia	Análise de mudança para identificar e caracterizar as causas da volatilidade dos requisitos.	Apresenta uma taxonomia que pode ser usada como uma abordagem para classificar mudanças de requisitos e identificar as causas dessas mudanças. A taxonomia ajuda a melhorar o conhecimento e o

ID	Contexto	Estratégia	Desafio	Resultado ou contribuição
				entendimento da volatilidade dos requisitos.
[RAM12]	<i>Sources of uncertainty at the requirement level</i>	Taxonomia	Vocabulário comum sobre fontes de incerteza a partir da perspectiva de um Sistema Adaptativo Dinâmico (SAD) e processo de análise, comparação e integração de diferentes técnicas e abordagens para lidar com a incerteza em um SAD.	Sintetizando conceitos de incerteza de outras disciplinas, este artigo explica o conceito de incerteza a partir da perspectiva de um SAD, propõe uma taxonomia de potenciais fontes de incerteza nas fases de requisitos, design e execução e identifica técnicas existentes para tipos específicos de mitigação da incerteza.
[ZAV97]	<i>Research efforts in Requirements Engineering</i>	Esquema de Classificação	A área da engenharia de requisitos é abrangente, interdisciplinar e totalmente aberta. Por estas razões, é difícil comparar, estender e explorar os resultados de pesquisas nesta área.	Este artigo apresenta um esquema de classificação para os esforços de pesquisa em engenharia de requisitos. O esquema de classificação ajuda: (1) delimitar a área e incentivar a cobertura de toda a área de pesquisa; (2) fornecer estrutura para incentivar a descoberta e a articulação de novos princípios, e (3) ajudar a agrupar coisas semelhantes, para comparar, estender, e explorar os resultados.

#### 2.4.7.2.5 Ferramentas e Métodos de Engenharia de Software

Para o SWEBOK [ABR04], as ferramentas de ES são as ferramentas baseadas em computador que se destinam a apoiar os processos de ciclo de vida do software. As ferramentas são muitas vezes projetadas para suportar determinados métodos de engenharia de software, reduzindo qualquer carga administrativa associada à aplicação do método manualmente. Os métodos de ES, segundo SWEBOK, têm como objetivo tornar as atividades de ES sistemáticas e com mais chances de ser bem sucedido.

Métodos geralmente oferecem uma notação e um vocabulário comum, procedimentos para a execução de uma determinada tarefa e diretrizes para a verificação tanto do processo como do produto final. A área de Ferramentas e Métodos abrange seis artigos sobre sistematização de ferramentas na área de ES. A Tabela 12 resume os desafios e as contribuições de cada estudo desta área.

Tabela 12 - Sistematização em Ferramentas e Métodos de ES.

ID	Contexto	Estratégia	Desafio	Resultado ou contribuição
[ALI10]	<i>Computer-supported Critics Tools</i>	Taxonomia	Existem várias definições para <i>critics</i> na literatura, porém não existe nenhuma categorização aceita sobre <i>critics tools</i> e também não existe uma forma sistemática de comparar e contrastar as diferentes <i>critics tools</i> .	Uma nova taxonomia para caracterizar os sistemas de crítica. A taxonomia contribui com uma visão geral da pesquisa sobre <i>critics tools</i> , ajuda a identificar e distinguir elementos-chave e a reconhecer técnicas ou métodos aplicados a <i>critics</i> .
[DAV88]	<i>Tools and techniques in Software Development Life cycle</i>	Taxonomia	Inconsistência da terminologia utilizada na indústria de software, por exemplo, os termos "requisitos", "especificação", ou "design". Dificuldade para categorizar as ferramentas e técnicas em desenvolvimento de software.	A taxonomia proposta pode ser usada para (1) ajudar a entender as relações entre as ferramentas e técnicas de requisitos existentes, (2) ajudar os usuários a determinar quais ferramentas de requisitos e técnicas que eles realmente precisam e o que está sendo oferecido a eles, e (3) permitir profissionais e pesquisadores que estão apresentando novas ferramentas e técnicas para categorizar e caracterizar seus produtos, em comparação com os outros.
[DEL04]	<i>Runtime Software-Fault Monitoring Tools</i>	Taxonomia	A diversidade das ferramentas de monitoramento em tempo de execução torna complexa a tarefa de identificar os tipos de	O estudo apresenta uma taxonomia na qual desenvolvedores e pesquisadores podem analisar e diferenciar as diversas abordagens de ferramentas de monitoramento

ID	Contexto	Estratégia	Desafio	Resultado ou contribuição
			aplicações que são suportados por um determinado monitor ou a identificação de propriedades do software.	em tempo de execução. A taxonomia também fornece uma linguagem comum para discutir sistemas de monitoramento em tempo de execução e seus componentes.
[NOV10]	<i>Static Code Analysis Tool</i>	Taxonomia	Diversidade grande de ferramentas para análise de código estático e dificuldade em comparar diversas categorias como: tecnologia, disponibilidade de regras, idiomas suportados, extensibilidade, etc.	Uma taxonomia para a classificação dos analisadores de código estático para ajudar na avaliação e classificação sistemática dessas ferramentas.
[SUL04]	<i>Simulation tools</i>	Taxonomia	Dificuldade em categorizar e analisar várias ferramentas de simulação para sistemas paralelos e distribuídos.	A taxonomia proposta prevê a classificação das ferramentas de simulação em categorias com base em diferentes aspectos e propriedades.
[UTT12]	<i>Model-based Testing Tools and Techniques classification</i>	Taxonomia	A popularidade do campo de teste baseado em modelos gerou um grande número de publicações, porém não existe uma forma unificada de comparação das diferentes abordagens de teste baseado em modelos.	Este artigo define uma taxonomia que ajuda a compreender as características, semelhanças e diferenças entre as diversas abordagens de teste baseado em modelo e fornece um framework para comparar e avaliar qualitativamente as ferramentas e técnicas de MBT.

Todos os estudos propuseram a criação de uma taxonomia para resolver os desafios apresentados e a principal motivação relatada foi que taxonomias ajudam a categorizar ou classificar as diversas ferramentas existentes e também promovem uma linguagem comum aos profissionais e pesquisadores para analisar e comparar as mesmas.

#### 2.4.7.2.6 Qualidade de Software

A área de conhecimento em Qualidade de Software compreende o conjunto de atividades relacionadas com garantia de qualidade de software e define atividades importantes como verificação e validação de software [SEW04].

Nas últimas duas décadas, a inspeção de software surgiu como uma das técnicas mais eficazes de garantia da qualidade em Engenharia de Software [LAI00]. Segundo [LAI00], a maioria dos trabalhos publicados não foi integrada em um contexto mais amplo, isto é, dentro de um corpo de conhecimento, tornando difícil conciliar e avaliar os trabalhos em uma fase específica do ciclo de vida do software. O autor complementa que esta pode ser uma razão pela qual os profissionais ainda estão fazendo perguntas do tipo: Quais são as principais diferenças entre as abordagens de inspeção disponíveis atualmente? Em que parte do ciclo de vida essas abordagens podem ser aplicadas? Que tipos de ferramentas de apoio estão disponíveis para inspeção de software? A Tabela 13 resume os desafios e as contribuições de cada estudo desta área.

Tabela 13 - Sistematização em Qualidade de Software.

ID	Contexto	Estratégia	Desafio	Resultado ou contribuição
[LAI00]	<i>Software Inspections</i>	Taxonomia	A área de inspeção de software cresceu muito e atingiu certa maturidade, no entanto, ainda não existe uma visão abrangente e sistemática desta área.	O estudo apresenta uma taxonomia que pode ajudar os profissionais a identificar experiências de inspeção diretamente relacionada a um estágio do ciclo de vida do desenvolvimento de software, estruturar a grande quantidade de estudos sobre inspeção e ajudar os pesquisadores a comparar e avaliar métodos de inspeção existentes.
[MAN03]	<i>Bad code smells</i>	Taxonomia	Existem muitos padrões de <i>Bad smells</i> e a comparação de relação e semelhança entre eles é complexa.	O estudo apresenta uma taxonomia que classifica os <i>bad smells</i> semelhantes. A taxonomia ajuda na compreensão dos tipos de <i>bad smells</i> e reconhece as relações entre eles.

#### 2.4.8 Conclusões do mapeamento sistemático

Tanto os dados da análise quantitativa como da análise qualitativa apresentaram resultados importantes a respeito da sistematização do conhecimento em ES. O mapeamento sistemático foi conduzido por meio de um protocolo de revisão que especificou os métodos utilizados durante a condução do trabalho. Os critérios definidos no protocolo foram necessários e suficientes para se obter os estudos primários necessários à realização da pesquisa. Esta seção tem por objetivo descrever as conclusões do estudo de mapeamento sistemático, com base no tipo de análise conduzida.

**Conclusão 1:** Existe um aumento crescente por sistematização do conhecimento em ES a partir da década de 2000.

O resultado do mapeamento sistemático selecionou mais de 90% dos estudos a partir da década de 2000, ou seja, 40 dos 43 artigos selecionados foram publicados a partir do ano de 2000. Apesar do número de estudos sobre sistematização em ES ter aumentado consideravelmente a partir da década de 2000, também encontramos estudos nas décadas de 80 e 90 que tinham como principal objetivo estudar formas de sistematizar o conhecimento desta área. Ainda sobre o gráfico da Figura 7, destaca-se o número de estudos encontrados na década de 2010, onde o resultado é maior que a década de 2000 com apenas três anos e meio (de 2010 a Junho de 2013).

O estudo mais antigo selecionado nesta revisão foi de 1988 [DAV88], no qual os autores abordaram as dificuldades para categorizar as ferramentas e técnicas em desenvolvimento de software durante o estágio inicial de desenvolvimento de software. Os autores descrevem problemas como a inconsistência da terminologia utilizada na indústria de software, por exemplo, os termos "requisitos", "especificação", ou "design". Este estudo apresentou uma taxonomia de quatro níveis para representar o ciclo de vida inicial do software. Os níveis são (1) análise das necessidades do usuário, (2) definição da solução, (3) definição do comportamento externo do software e (4) design preliminar. O objetivo principal da taxonomia deste artigo é fornecer para indústria uma forma sistemática para compreensão das atividades distintas que ocorrem nas fases iniciais do desenvolvimento de software e também para que diferentes metodologias possam ser comparadas.

Apenas dois estudos foram selecionados na década de 90, sendo que o estudo [PRE94] tinha como objetivo compreender as diferenças e semelhanças entre diferentes

áreas da Ciência da Computação e o segundo, [ZAV97], descreve a dificuldade inerente ao processo de comparar, estender e explorar os resultados de estudos na área de Engenharia de Requisitos, por ser uma área abrangente, interdisciplinar e totalmente aberta. Segundo [PRE94], o foco da área de ES é em compreender, controlar, gerenciar e melhorar processos e produtos de software baseado nos princípios da engenharia. Por outro lado, na área de Interação Humano-Computador, o foco está em entender a usabilidade dos produtos de softwares, descrever critérios para boas interfaces humano-computador e encontrar métodos e ferramentas para desenhar e desenvolver essas interfaces. Os autores identificaram nesse estudo que existem semelhanças e diferenças nas técnicas de medição utilizadas nas duas áreas e apresentaram uma taxonomia para explicar o processo de medição, examinar pesquisas anteriores em ambos os campos e também para facilitar a comunicação entre as duas disciplinas, fornecendo um vocabulário comum. A taxonomia descreve quatro dimensões: (1) objetivo do estudo, (2) plano do estudo, (3) métodos do estudo e (4) técnicas do estudo. Cada dimensão foi subdividida em níveis secundários para facilitar a caracterização dos estudos. Por outro lado, o estudo [ZAV97] apresenta um esquema de classificação para as pesquisas em Engenharia de Requisitos (ER). Segundo os autores, o esquema de classificação ajuda a (1) delimitar a área e incentivar a cobertura de toda a área de pesquisa, (2) fornecer uma estrutura para incentivar a descoberta e a articulação de novos princípios, e (3) ajudar a agrupar coisas semelhantes, para comparar, estender, e explorar os resultados dos estudos nessa área.

**Conclusão 2:** Existe uma forte tendência por estudos sobre sistematização do conhecimento na área de Processo de Engenharia de Software do SWEBOK.

Aproximadamente 40% dos estudos selecionados (Figura 9), apresentaram desafios e estratégias para sistematização do conhecimento na área de Processo de Engenharia de Software. Forward e Lethbridge [FOR08] descreve que a pesquisa empírica de software poderia ser melhorada se houvesse uma maneira sistemática para identificar os tipos de software para o qual se aplica a evidência empírica e propõe uma taxonomia que pode ser usada para ajudar a categorizar estudos empíricos e outras provas obtidas por pesquisadores de ES, de modo que a aplicabilidade da evidência possa ser entendida de forma mais sistemática. O autor complementa que muitas vezes os resultados dos estudos não são aplicáveis em um cenário global, mas é mais provável que se apliquem apenas em certos contextos, tais como o tipo de software em que a evidência foi testada e adicionam que a taxonomia também pode ser utilizada na

aplicação prática dos resultados da pesquisa. Por exemplo, os resultados da pesquisa de uma categoria de software podem ser aplicáveis em categorias semelhantes dentro da taxonomia.

**Conclusão 3:** A criação de taxonomias para sistematização do conhecimento mostrou-se como a estratégia mais utilizada pelos pesquisadores.

Entre as estratégias de sistematização pesquisadas neste trabalho, a taxonomia foi a mais utilizada nos estudos, totalizando 31 dos 43 artigos selecionados na revisão. O principal objetivo dos pesquisadores era classificar um determinado tópico de pesquisa da ES. Os esquemas de classificação também são uma alternativa para o processo de sistematização do conhecimento, uma vez que estes artefatos ajudam os pesquisadores em diversas tarefas de sistematização, tais como, realizar um mapeamento sistemático de uma área [MEHM13 e PET11], melhorar o processo de descrição do contexto dos estudos [SMI08 e OLU07] e ajudar a classificar uma determinada área da ES [SAK09 e ZAV97]. A união de duas estratégias de sistematização também é possível, conforme mostrou os artigos [FRE13, SMI12 e VAS11], nos quais apresentaram uma taxonomia e uma terminologia unificada como estratégias para sistematização do conhecimento.

Smite et al. [SMI12] descreve que a diversidade no termo *outsourcing*, por exemplo, tem causado dificuldades em determinar qual termo usar em que situação e dessa forma acabam criando novos obstáculos para buscar e encontrar estudos relevantes durante uma revisão sistemática da literatura. Assim, a necessidade de criar uma terminologia precisa, com definições para as diferentes situações de *sourcing* global surge como um caminho para a comunidade desenvolver novos trabalhos com base em outros estudos e, portanto, progredir mais rapidamente. Os autores apresentam uma proposta de taxonomia para classificar e mapear o conhecimento existente em DDS e descrevem que o uso da taxonomia pode ajudar pesquisadores e profissionais na compreensão dos artigos que documentam descobertas e experiências em DDS. Os autores destacam que a taxonomia tem dois usos primários. Primeiro, fornece uma base para os pesquisadores classificarem seus próprios estudos e estudos relacionados. Em segundo lugar, uma vez que os estudos são classificados de acordo com a taxonomia, é possível identificar um conjunto de estudos que descrevem uma situação particular. Os pesquisadores podem usar a taxonomia para sintetizar o conhecimento existente, identificar lacunas na literatura e também encontrar trabalhos relacionados, assim como,

os profissionais podem usá-la para encontrar evidências de uma pergunta específica em um contexto que é semelhante ao seu.

O estudo [SMI08] apresenta um conjunto de atributos que são necessários para avaliar a aplicabilidade e capacidade de generalização dos resultados empíricos em DDS e juntos, esses fatores constituem um esquema de classificação. Os autores agruparam os fatores em três categorias principais: (1) histórico em DDS, ou seja, fatores que diferenciam os cenários em DDS e classificam a origem das conclusões relatadas, (2) histórico empírico dos fatores que se referem à origem e à perspectiva dos dados empíricos utilizados para obter as conclusões relatadas e (3) histórico do estudo, através das características e aspectos dos estudos investigados. Por fim os autores descrevem que um esquema de classificação pode ser usado para compreender o contexto dos estudos existentes e orientar o processo de informação e estruturação de estudos futuros nesta área.

#### 2.4.9 Limitações do mapeamento sistemático

Este mapeamento sistemático possui algumas limitações. A primeira delas está relacionada ao fato de que os estudos selecionados foram todos da biblioteca online SCOPUS. Esta base de dados foi escolhida, pois indexa publicações de diversas bases e também por fornecer recursos para extração de dados de forma eficiente. Porém, é possível que outras bases não utilizadas nesse trabalho também contenham artigos sobre sistematização do conhecimento em ES.

Outra limitação desse trabalho está relacionada ao viés dos pesquisadores durante o processo de análise dos artigos. Para minimizar esta limitação, a revisão foi realizada por dois pesquisadores, conforme recomendações na literatura específica da área. O pesquisador com mais experiência validou as etapas do mapeamento sistemático como a classificação dos artigos, a extração dos dados e a síntese dos resultados provendo um *feedback* detalhado sobre os artefatos que foram desenvolvidos.

### 3 METODOLOGIA DE PESQUISA

Neste capítulo apresenta-se a metodologia de pesquisa utilizada no estudo. A seção 3.1 apresenta o desenho de pesquisa. Na seção 3.2 identificam-se os aspectos metodológicos do estudo.

A fim de compreender como podemos sistematizar as evidências empíricas geradas a partir das pesquisas científicas conduzidas na área de DDS optou-se pela realização de um estudo empírico, de natureza exploratória. Com relação à natureza do estudo, a pesquisa exploratória, que muitas vezes constitui-se na primeira etapa de uma investigação mais ampla, tem como principal finalidade desenvolver, esclarecer e modificar conceitos e ideias, com vistas à formulação de novas teorias, modelos e hipóteses [GIL10]. Gil [GIL10] acrescenta que as pesquisas exploratórias têm o objetivo de proporcionar uma visão geral, do tipo aproximativo, acerca de determinado fato, principalmente quando o tema escolhido é pouco explorado, dificultado a formulação de hipóteses específicas. Segundo Yin [YIN10], o estudo de caso exploratório auxilia em áreas onde há pouca literatura sobre um determinado tópico, então uma instância do contexto real é investigada, auxiliando na identificação de tópicos para pesquisa.

Assim, esta pesquisa se caracteriza como um estudo exploratório, e foram utilizados métodos primários (*survey* com especialistas) e secundários de pesquisa (revisão sistemática da literatura, do tipo mapeamento sistemático). Primeiramente realizou-se um estudo de mapeamento sistemático com o objetivo de identificar (1) quais são as áreas da Engenharia de Software que apresentam estudos sobre sistematização o conhecimento, (2) quais são os principais desafios relatados e (3) quais foram os resultados obtidos. O método de pesquisa utilizado nos estudos primários foi uma *survey* com especialistas, também conhecido como *Expert Opinion Survey* [FIN03, KIT08], que teve como objetivo identificar que tipo de informação de contexto é relevante, para pesquisadores e profissionais da indústria, em estudos com evidências empíricas e que muitas vezes não estão claras. Como instrumento de pesquisa utilizou-se um roteiro para uma entrevista semiestruturada, com questões abertas e fechadas.

Assim como muitas pesquisas, esta também apresenta suas limitações, principalmente no que se refere ao número de entrevistas realizadas, restringindo a generalização dos resultados obtidos. No Capítulo 6 deste estudo, aborda-se com mais profundidade a questão das limitações da pesquisa.

### 3.1 DESENHO DE PESQUISA

O desenho de pesquisa contempla as etapas necessárias para se alcançar o objetivo do estudo. Embora o desenho de pesquisa apresentado no PEP (Plano de Estudo e Pesquisa) contemple duas etapas, o escopo desta dissertação de mestrado foi replanejado no seminário de andamento e concentra-se, por limitações de tempo, nas atividades da etapa um. A seguir apresenta-se o desenho de pesquisa na Figura 12, constituído por quatro fases na primeira etapa, descritas a seguir:

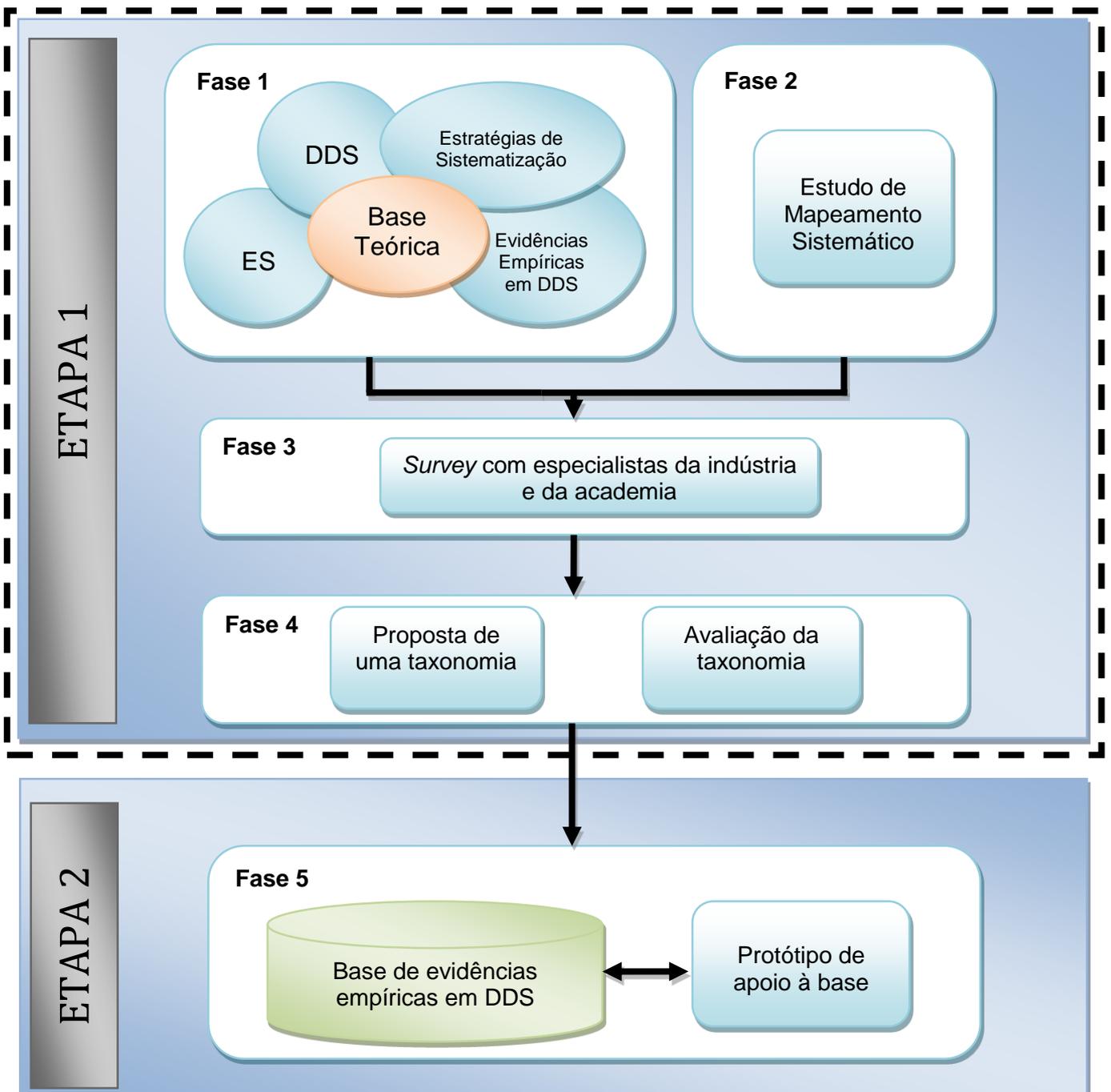


Figura 12 - Desenho de pesquisa.

A seguir são descritas as quatro fases contempladas nessa pesquisa:

- Fase 1: foi estudada a base teórica da área. Teve por objetivo complementar o conhecimento teórico a respeito da Engenharia de Software (ES), o Desenvolvimento Distribuído de Software (DDS), estratégias de sistematização, e evidências empíricas geradas a partir de estudos científicos da área de DDS. A estratégia adotada nesta fase foi à análise de dois artigos científicos relacionados com este tema, colaborando na identificação da questão de pesquisa deste trabalho;
- Fase 2: uma vez que a questão de pesquisa foi desenvolvida, planejou-se a segunda fase que consistia na execução de um estudo secundário, um mapeamento sistemático, abrangendo todas as áreas da ES. O objetivo foi compreender, de forma mais ampla, como estratégias de sistematização, tais como, taxonomias, esquemas de classificação e terminologias unificadas, poderiam contribuir para a sistematização do conhecimento em ES. Através da análise crítica dos resultados do estudo de mapeamento sistemático, foi possível categorizar os estudos nas estratégias de sistematização definidas e identificar desafios e soluções de diversas áreas da ES, incluindo a área de DDS;
- Fase 3: foi realizada uma *survey* com especialistas com o objetivo de mapear, na academia e na indústria, que tipo de informações de contexto profissionais e pesquisadores desta área procuram em estudos com evidência empírica em DDS. A população envolveu profissionais experientes que trabalham e pesquisam na área de DDS. Essa fase envolveu entrevistas com profissionais de diferentes empresas e instituições de ensino ao redor do mundo. No PEP foi proposta a execução de uma *survey*, através da aplicação de um questionário online semiestruturado com questões abertas e fechadas. No entanto, devido à complexidade das questões para a coleta de dados desta pesquisa, decidiu-se, durante a interação com a banca de avaliação do seminário de andamento, que o método de coleta de dados mais adequado seria através de entrevistas com os especialistas da área;
- Fase 4: esta fase contemplou duas atividades: a proposição de uma taxonomia e a avaliação da mesma. A primeira atividade foi à proposta de uma taxonomia para sistematização das evidências empíricas em DDS e a segunda atividade teve como objetivo classificar um conjunto de estudos

empíricos com o objetivo de fazer uma avaliação inicial da taxonomia proposta.

Conforme dito anteriormente, a etapa dois ficou de fora do escopo desta dissertação de mestrado, sendo parte de estudos futuros. O objetivo da etapa dois seria construir uma base de evidências empíricas em DDS e uma ferramenta de apoio a esta base. A ferramenta de suporte seria desenvolvida para formulação de consultas a base e também serviria de apoio à criação de novas taxonomias. A base de evidências empíricas permitiria a criação de uma terminologia mais completa em DDS e principalmente, a categorização das evidências empíricas nesta área.

O estudo de base teórica, contemplado na fase 1, foi importante na medida em que formou uma base teórica consistente para a continuidade do estudo, ao complementar o referencial teórico e compreender os principais desafios da área de DDS, assim como, conceitos e estratégias de sistematização das evidências empíricas desta área. Os artigos analisados nessa etapa foram importantes para definição do estudo secundário executado na segunda fase desta pesquisa.

O estudo de mapeamento sistemático, realizado na fase 2, analisou os desafios e soluções propostas para sistematização do conhecimento em ES. O mapeamento sistemático foi importante para consolidar a questão de pesquisa deste estudo e aprofundar o conhecimento sobre estratégias de sistematização utilizadas nos estudos selecionados na revisão. Esta revisão permitiu a identificação dos trabalhos já existentes sobre o tema e a extração de resultados e conclusões importantes a respeito da sistematização do conhecimento em ES. A análise dos resultados do mapeamento sistemático, assim como, o aprofundamento da base teórica realizado na fase 1, foram importantes para a construção do roteiro de entrevista aplicado na *survey* com especialistas (fase 3). O Apêndice A apresenta o protocolo do mapeamento sistemático e o Apêndice B lista os estudos resultantes desta revisão.

A *survey* com especialistas permitiu o estudo aprofundado sobre a sistematização das evidências empíricas em DDS sob a ótica da academia e da indústria, permitindo uma análise detalhada sobre os principais desafios relatados e informações de contexto dos estudos que são relevantes para pesquisadores e profissionais da indústria. Dessa forma, optou-se por empregar a *survey* na fase 3, utilizando como instrumento de coleta de dados um roteiro para entrevista semiestruturada, com questões abertas e fechadas. O Apêndice C apresenta o protocolo de estudo desenvolvido para suportar a geração, aplicação e análise dos resultados da *survey*. A definição e a utilização de protocolos para

desenvolvimento e formalização da *survey*, que se baseiam em um consistente referencial teórico, tiveram por objetivo uniformizar e sistematizar a tarefa de análise, aumentando a confiabilidade do estudo [PRI03].

A fase 4 envolveu a criação e a avaliação de uma taxonomia, contemplando as informações obtidas da base teórica e a consolidação dos dados extraídos da *survey*. Com o objetivo de ilustrar e avaliar a taxonomia proposta classificou-se um conjunto de estudos empíricos extraídos da principal conferência internacional em DDS: a *IEEE International Conference on Global Software Engineering (ICGSE<sup>4</sup>)*. Como resultado desta avaliação, os artigos selecionados foram classificados de acordo com quatro cenários que serão detalhados no Capítulo 5 deste trabalho.

### 3.2 ASPECTOS METODOLÓGICOS

De forma a garantir a confiabilidade deste estudo, um rigoroso processo de pesquisa foi planejado. Nesta pesquisa, a definição e utilização de protocolos para desenvolvimento e formalização do estudo tiveram por objetivo sistematizar a tarefa de análise, aumentando a confiabilidade da pesquisa. O mapeamento sistemático foi planejado utilizando um protocolo de acordo com Kitchenham e Charters [KIT07]. A *survey*, além de um protocolo formal, passou por uma validação de face e conteúdo e um pré-teste com o objetivo de garantir a integridade dos resultados. Os protocolos para o mapeamento sistemático e a *survey* encontram-se no Apêndice A e C desta dissertação, respectivamente.

O viés de todo o processo de pesquisa foi minimizado com a avaliação das etapas de pesquisa com um pesquisador experiente (o professor orientador). No mapeamento sistemático, a avaliação ocorreu principalmente no protocolo, na ficha de leitura dos artigos e na análise dos resultados. No estudo de caso, o protocolo da coleta dos dados e a análise dos resultados também contaram com esta avaliação.

---

<sup>4</sup> <http://www.icgse.org>

## 4 SURVEY

Este capítulo apresenta a *survey* desenvolvida. Na seção 4.1 apresenta-se a base metodológica da *survey*, na seção 4.2 são vistas as fases e a operacionalização da *survey* e os resultados obtidos da academia e da indústria são apresentados na seção 4.3. Na seção 4.4 descreve-se a consolidação dos resultados e por fim, na seção 4.5, apresentam-se as lições extraídas da *survey* que serviram de base para a criação dos elementos da taxonomia proposta.

### 4.1 BASE METODOLÓGICA

Neste estudo, planejou-se e executou-se uma *survey* com base na opinião de especialistas [FIN03, KIT08]. Apesar de a *survey* ser um método de pesquisa popular, ela deve ser utilizada em condições apropriadas. De acordo com Fink [FIN03], a *survey* não é apenas um questionário ou uma lista de verificação para a coleta de informações, e sim um método de pesquisa utilizado para coletar dados para descrever, comparar ou explicar conhecimentos, atitudes e comportamentos. O método *survey* permite a generalização sobre as crenças e opiniões de muitas pessoas, estudando apenas um subconjunto delas [KAS05].

A *survey* pode ser confundida com o estudo de caso, no entanto ao analisar o conceito de YIN [YIN05], de que o estudo de caso é utilizado para responder perguntas “como” e “porque”, entendeu-se que não é nenhuma destas perguntas que se pretende responder. Além disso, o estudo de caso acaba sendo descartado como método de pesquisa, pois nesta situação não se tem um caso a analisar e sim um questionamento a ser explorado, a fim de ter como saída uma lista de atributos que serão utilizados para a criação de uma taxonomia. A pergunta que será respondida nesta etapa da pesquisa é “quais evidências empíricas e informações de contexto, em estudos sobre Desenvolvimento Distribuído de Software (DDS), são relevantes para profissionais da indústria e academia?”. Desta forma, a *survey* é o método de pesquisa que identificou-se como o mais adequado para responder os questionamentos deste estudo.

A *survey* realizada teve como base a coleta de dados qualitativos. Ele foi realizado com profissionais de várias organizações e instituições de ensino que possuíam experiência com DDS. O objetivo da *survey* foi identificar que tipo de informação de contexto é relevante para profissionais e pesquisadores desta área e serviu para, juntamente com as informações coletadas no mapeamento sistemático, elaborar uma

taxonomia para sistematização das evidências empíricas geradas a partir de estudos científicos em DDS.

#### 4.1.1 Seleção das organizações e unidade de análise

A unidade de análise do estudo foi definida como sendo profissionais que trabalham e pesquisam na área de Desenvolvimento Distribuído de Software. Desta forma, foram escolhidos, por conveniência, profissionais que já tinham tido alguma experiência em algum projeto com desenvolvimento distribuído ou que já tinham desenvolvido projetos de pesquisa nesta área.

Os profissionais selecionados pertenciam a dez organizações diferentes, sendo quatro empresas multinacionais e seis universidades. Todas colaboraram no processo de coleta de dados por meio das entrevistas.

#### 4.1.2 Fonte de dados e seleção de participantes

A coleta de dados foi constituída por fontes primárias. As fontes primárias foram constituídas de entrevistas. Foram realizadas entrevistas semiestruturadas com questões abertas e fechadas. Partiu-se de um roteiro com questões formuladas aos entrevistados e adequadas conforme seu desenvolvimento.

O critério inicial para a definição dos entrevistados centrou-se na unidade de análise e nos objetivos do estudo. Neste sentido, a população envolvida constituía-se de profissionais da indústria que possuíam o perfil de executivo, gerente, líder de projeto ou membro do time de projeto e pesquisadores com experiência na área de DDS. O instrumento de coleta de dados consistiu-se de um roteiro para entrevista semiestruturada (Apêndice C). As entrevistas foram organizadas para identificar atributos de contexto que são relevantes para os profissionais entrevistados, para que posteriormente, possam ser usados para propor uma taxonomia em DDS.

#### 4.1.3 Análise de Dados

A técnica utilizada para a análise de dados foi a de análise de conteúdo [MOZ10]. Desta forma, todas as entrevistas foram gravadas, transcritas e analisadas posteriormente. Com o objetivo de familiarizar o pesquisador com os dados coletados antes de iniciar a codificação destes, realizou-se uma leitura detalhada das transcrições. Para cada uma das questões foram identificadas categorias preliminares, para as quais os dados foram codificados. Este processo foi desenvolvido independentemente pelo

pesquisador e após consolidado com o orientador, definindo um conjunto definitivo de categorias a serem consideradas.

#### 4.2 FASES E OPERACIONALIZAÇÃO DA SURVEY

A *survey* contou com a participação de 15 profissionais especialistas em DDS, sendo sete respondentes da academia e oito respondentes da indústria. Em um primeiro momento, buscaram-se indicações com o professor orientador de profissionais e pesquisadores experientes na área de DDS. Após entrar em contato com cada um dos indicados, os participantes dispuseram cerca de uma hora de tempo para as entrevistas que ocorreu tanto de forma presencial quanto a distância (via *Skype*<sup>5</sup>), quando os participantes estavam fisicamente distantes em outras cidades ou outros países. Ao todo foram seis entrevistas presenciais e seis através da ferramenta *Skype*. Apenas três respondentes enviaram as respostas por e-mail devido à falta de disponibilidade.

Um questionário semiestruturado foi utilizado como instrumento de coleta. Este questionário foi desenvolvido tomando-se por base um roteiro inicial de questões, a partir do aprofundamento da base teórica e dos resultados do estudo de mapeamento sistemático executado e foi representado pelo protocolo de pesquisa desenvolvido para a *survey* (Apêndice C).

A validação de face e conteúdo do protocolo de pesquisa foi realizada por uma pesquisadora sênior (doutora) da *University of Maryland Baltimore County (UMBC)*<sup>6</sup>, nos Estados Unidos. A partir disso foi executado um pré-teste, com um pesquisador que é aluno do doutorado da PUCRS. Após sua aplicação foi possível descobrir os inconvenientes, eliminar equívocos e ambiguidades e escolher a formulação mais adequada das perguntas para a finalidade da pesquisa. Foram realizadas iterações sucessivas até gerar uma versão estável do roteiro. Uma vez que parte das entrevistas seria realizada em inglês e parte em português, adotou-se como uma prática necessária a validação de face e conteúdo por pesquisadores nos dois idiomas. Da mesma forma, o roteiro de entrevista foi transcrito nos dois idiomas (inglês e português).

Após a transcrição das entrevistas, desenvolveu-se a análise qualitativa dos dados coletados na *survey*. Foi realizada uma análise de conteúdo, na qual foram identificadas as categorias preliminares, passando pela definição do universo estudado e a definição destas categorias, sendo este o procedimento mais importante, visto que faz a conexão

---

<sup>5</sup> <http://www.skype.com>

<sup>6</sup> <http://www.umcp.umd.edu>

entre os objetivos de pesquisa e os resultados. Seu valor fica sujeito à legitimidade das categorias de análise e depende da qualidade da elaboração conceitual feita *a priori* pelo pesquisador e da exatidão com a qual ele consiga traduzir os textos em categorias, permitindo, desta forma, formular conclusões e obter novas informações por meio do exame detalhado dos dados [PRI03].

#### 4.2.1 Perfil da população

A Tabela 14 sintetiza o perfil de cada respondente, descrevendo a ocupação principal, país onde os profissionais trabalham, função, anos de experiência com DDS na indústria e na academia e o grau de escolaridade.

Tabela 14 - Perfil dos respondentes da *survey*.

ID	Ocupação	País	Função	Experiência	Experiência	Educação
				Indústria	Academia	
1	Academia	Suécia	Prof / Pesq	NA	De 6 a 10 anos	Pós-doutorado
2	Academia	Itália	Prof / Pesq	NA	Mais de 10 anos	Doutorado
3	Academia	Itália	Prof / Pesq	De 2 a 4 anos	De 6 a 10 anos	Pós-doutorado
4	Academia	Alemanha	Prof / Pesq	De 4 a 6 anos	De 4 a 6 anos	Pós-doutorado
5	Academia	Brasil	Prof / Pesq	De 2 a 4 anos	Mais de 10 anos	Doutorado
6	Academia	Brasil	Prof / Pesq	De 2 a 4 anos	De 2 a 4 anos	Doutorado
7	Academia	Brasil	Prof / Pesq	NA	Mais de 10 anos	Doutorado
8	Indústria	Brasil	Gerente	Mais de 10 anos	De 2 a 4 anos	Mestrado
9	Indústria	Brasil	Gerente	Mais de 10 anos	Mais de 10 anos	Mestrado
10	Indústria	Brasil	Gerente	De 6 a 10 anos	De 2 a 4 anos	Mestrado
11	Indústria	Brasil	Líder de Projeto	Mais de 10 anos	Menos de 2 anos	Mestrado
12	Indústria	Brasil	Líder de Projeto	De 6 a 10 anos	De 2 a 4 anos	Mestrado
13	Indústria	Índia	Executivo	Mais de 10 anos	Mais de 10 anos	MBA
14	Indústria	EUA	Pesquisador	De 6 a 10 anos	Mais de 10 anos	Pós-doutorado
15	Indústria	Índia	Pesquisador	De 2 a 4 anos	De 2 a 4 anos	MBA

As entrevistas contaram com a atenção dos participantes e foram fornecidas todas as informações solicitadas, sempre respeitando a política de privacidade e confidencialidade do participante.

#### 4.2.2 Roteiro de entrevista

O roteiro de entrevista constitui-se de um questionário semiestruturado com questões abertas e fechadas. As questões foram analisadas conforme quatro dimensões,

descritas no Apêndice C deste trabalho: motivações, expectativas, desafios e contexto dos estudos.

A primeira questão buscava compreender, de uma forma mais ampla, para qual finalidade os respondentes buscariam artigos empíricos na literatura relacionados com a área de pesquisa em DDS. Na segunda questão, buscaram-se explorar quais eram as expectativas dos respondentes com relação aos estudos empíricos em DDS, identificando, de uma forma mais ampla, o que os participantes esperavam encontrar nesses estudos.

Com relação aos desafios inerentes ao processo de busca por artigos empíricos em DDS, foram exploradas duas questões. A primeira questão buscou mapear na, indústria e na academia, como os profissionais que trabalham e pesquisam em DDS percebem a complexidade do processo de busca por artigos empíricos nesta área e na segunda questão foram explorados os motivos de cada resposta.

As questões da dimensão de contexto tiveram como objetivo identificar que tipo de informação de contexto dos estudos em DDS é relevante para profissionais e pesquisadores desta área. Neste estudo, as informações de contexto estão relacionadas com as características do estudo, tais como, método de pesquisa, coleta de dados, características da população estudada (número de participantes, perfil, etc.), características do objeto de estudo (empresas, projetos, ferramentas, etc.), nível de dispersão (mesma localização física, distância nacional, distância continental ou distância global), entre outros.

As questões de contexto foram ilustradas com dois estudos empíricos publicados na revista *Software Process Improvement and Practice*<sup>7</sup> em 2008. As fontes dos artigos foram selecionadas do estudo [SMI12]. Os critérios utilizados para escolha dos artigos foram: artigo de revista, estudo empírico e classificado na taxonomia proposta por [SMI12]. Primeiramente iniciou-se com uma questão fechada, onde os respondentes tinham que selecionar as informações de contexto que eram mais relevantes nos dois estudos empíricos selecionados. O objetivo desta questão era ilustrar quais informações de contexto eram mais relevantes para academia e indústria. Para fins de análise dos resultados, as sentenças de cada artigo foram identificadas com um padrão, ou seja, as sentenças do “Artigo A” foram identificadas como A1, A2 e assim sucessivamente até a última sentença A10 e da mesma forma, as sentenças do “Artigo B” com o padrão B1 até B10. Na segunda questão, buscou-se explorar, através de questões abertas, as razões

---

<sup>7</sup> [http://onlinelibrary.wiley.com/journal/10.1002/\(ISSN\)1099-1670](http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1099-1670)

pelas quais as sentenças selecionadas eram importantes para os respondentes e quais informações de contexto não estavam claras nos estudos analisados.

### 4.3 RESULTADOS DA SURVEY

Esta seção apresenta os resultados da *survey* com especialistas realizado. Os resultados foram analisados de forma separada para academia e para indústria e posteriormente foram consolidados. Portanto, na seção 4.3.1 são relatados os resultados da academia e na seção 4.3.2 os resultados da indústria são apresentados.

#### 4.3.1 Resultados da academia

A seguir serão apresentadas as caracterizações dos respondentes e os elementos de análise, assim como os resultados obtidos da academia.

##### 4.3.1.1 Caracterização dos respondentes da academia

Foram realizadas entrevistas com sete pesquisadores, sendo que deste total, seis foram realizadas através da ferramenta Skype e apenas uma entrevista foi presencial. As entrevistas realizadas com o Skype permitiram colher as informações de respondentes de outros estados do Brasil e também outros países, tais como, Suécia, Itália e Alemanha. As entrevistas tiveram uma duração média de 28 minutos (entre um mínimo de 21 minutos e um máximo de 42 minutos).

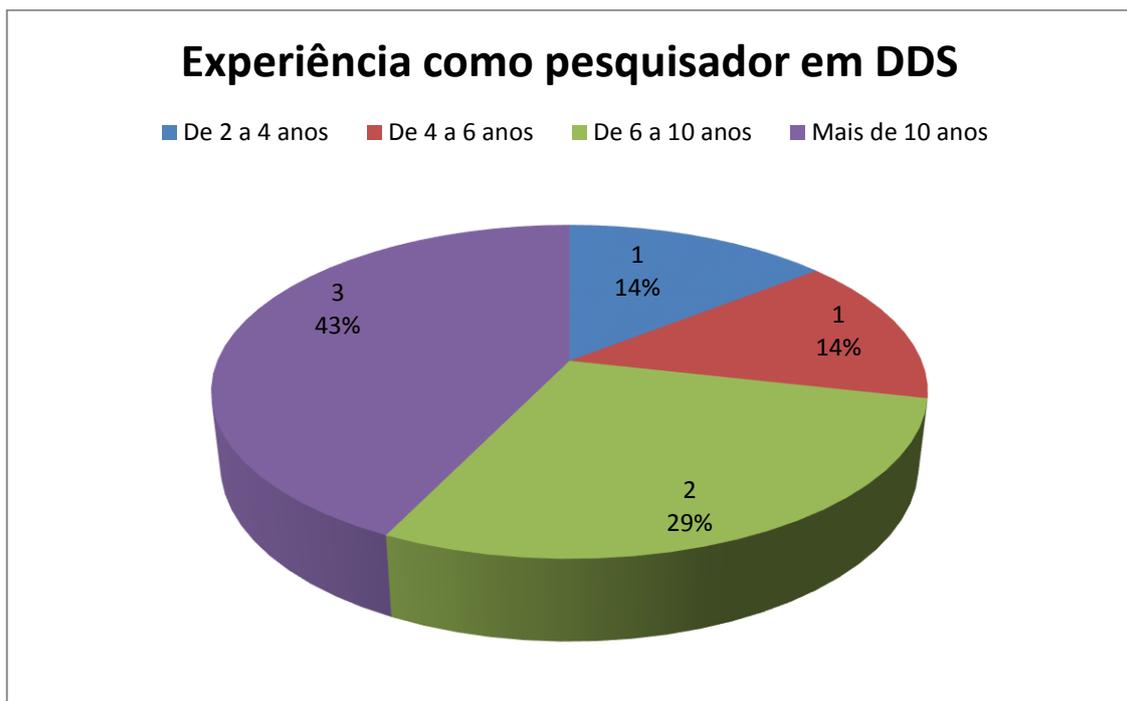


Figura 13 – Tempo de experiência como pesquisador em DDS.

Todos os entrevistados possuem experiência com pesquisa em DDS, sendo que dois possuem experiência entre 6 e 10 anos e 3 possuem mais de 10 anos (Figura 13). Dos sete respondentes, apenas três não possuem experiência na indústria, o que é aceitável, uma vez que os participantes são dedicados à academia e possuem em média mais de seis anos de experiência como pesquisador na área de DDS. Os demais entrevistados possuem no mínimo de 2 a 4 anos de experiência na indústria.

Com relação ao nível de formação, o grupo representa adequadamente o alto nível de qualificação, sendo que todos possuem no mínimo o título de doutor. Destes, três possuem pós-doutorado. Quanto à distribuição geográfica, a Figura 14, apresentada a seguir, mostra que três respondentes trabalham no Brasil e os demais no continente europeu.

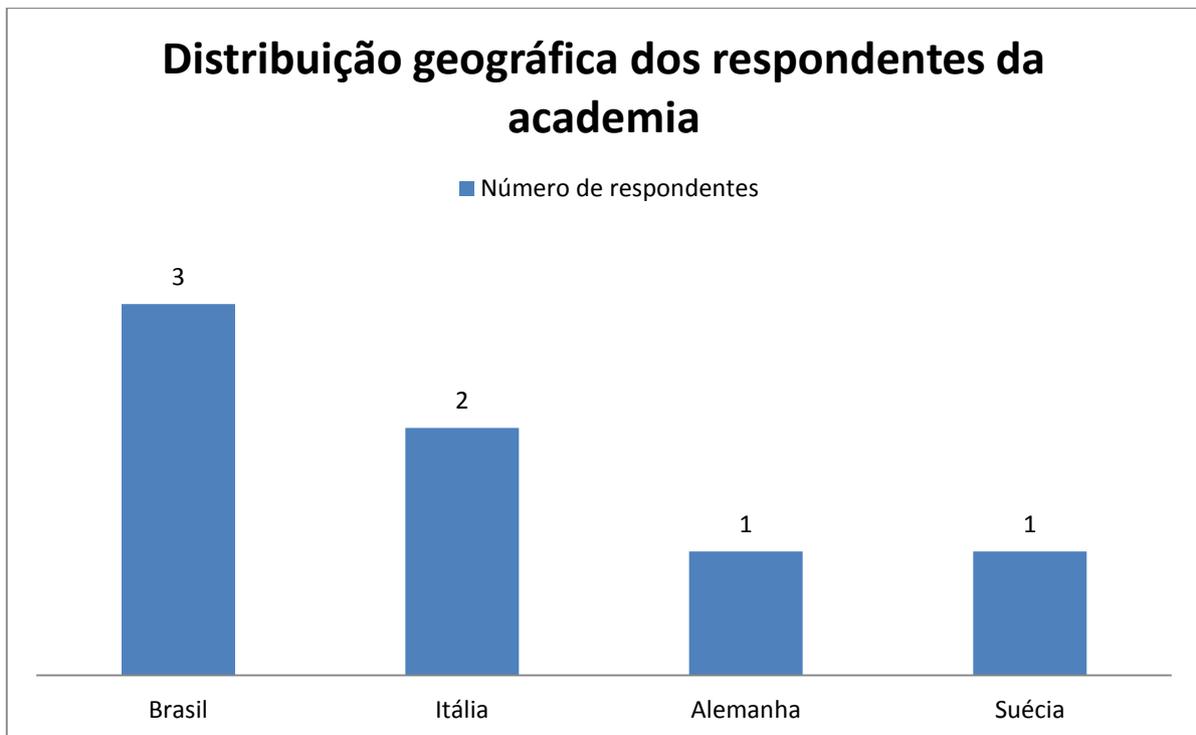


Figura 14 – Distribuição geográfica dos respondentes da academia.

#### 4.3.1.2 Elementos de análise da academia

A seguir, apresentam-se os elementos analisados e as categorias obtidas, buscando direcionar para um conjunto de lições relevantes visando contemplar o objetivo principal deste estudo, relacionado com a busca de uma taxonomia para sistematização das evidências empíricas em DDS. Esta análise permitiu relatar os resultados de forma a traduzir a realidade estudada e sua relação com os objetivos desta pesquisa.

#### 4.3.1.2.1 Motivações

Durante as entrevistas, os respondentes citaram diversas motivações pelas quais buscariam artigos empíricos descrevendo estudos científicos ou relatos de experiência de outras empresas, tais como, para responder as perguntas que são colocadas pelos parceiros da indústria, obter novas fontes de estudos relacionados com um determinado tema de pesquisa, encontrar soluções de problemas vindo da prática e buscar estudos que possam servir como fundamentação teórica para uma nova pesquisa.

Alguns trechos das entrevistas transcritas permitem ilustrar os resultados, como, por exemplo, este de um pesquisador da Suécia:

*“A principal motivação seria para obter as fontes de estudos relacionados sobre DDS. [...] E para responder as perguntas que são colocadas pelos nossos parceiros da indústria.”*

Assim como, o relato de um pesquisador do Brasil com mais de 10 anos de experiência com pesquisa nesta área:

*“Para pesquisa aplicada, ou seja, encontrar soluções de problemas vindas da prática, do mundo real. [...] Encontrar soluções para problemas práticos já descritos na literatura. Encontrar artigos científicos que sirvam como inspiração para solucionar problemas novos ou para ajudar a percorrer os caminhos que outros pesquisadores já utilizaram.”*

E o relato de outro pesquisador do Brasil que abordou a questão sobre a fundamentação teórica:

*“Buscar motivação para novas pesquisas na área. [...] Buscar elementos de comparação para ajudar a construir a pesquisa. Fundamentação teórica, ou seja, buscar estudos que possam servir como fundamentação para nossa pesquisa e, portanto, serem usados com objetivo de comparação e avaliação dos resultados.”*

Os pesquisadores também relataram outras finalidades, como, buscar relatos de dificuldades inerentes ao DDS e como elas foram resolvidas, conhecer o estado da arte da área e estudar abordagens comprovadas em DDS. Como ilustração, o relato de um pesquisador do Brasil sobre quais seriam as suas principais motivações para buscar estudos empíricos:

*“Buscar relatos de experiências e dificuldades inerentes ao DDS que foram descritos por empresas que estão inseridas nesse contexto e como as*

*dificuldades foram superadas. [...] Com base nessas experiências conseguir alavancar ou definir algum modelo ou metodologia para aplicar em alguma linha de pesquisa.”*

Alguns respondentes relataram temas específicos de pesquisa como comunicação, confiança, gerenciamento de projetos distribuídos e redução de custo com o DDS. A seguir, apresenta-se um trecho transcrito de um pesquisador da Itália:

*“Buscar conhecer novos métodos, guidelines e lições aprendidas sobre como gerenciar projetos distribuídos no contexto de DDS. [...] Encontrar estudos empíricos com teorias ou métodos com evidências claras do que funcionou e o que não funcionou no estudo em questão.”*

A seguir, apresenta-se na Tabela 15, a síntese das principais motivações relatadas pela academia em relação à busca por estudos empíricos em DDS.

Tabela 15 – Síntese das motivações da academia para busca de estudos empíricos.

Motivações da academia
Fundamentação teórica
Responder perguntas da indústria
Encontrar soluções para problemas práticos e dificuldades inerentes ao DDS
Conhecer o estado da arte da área

#### 4.3.1.2.2 Expectativas

Buscou-se com a segunda questão, explorar quais eram as expectativas dos respondentes com relação aos estudos empíricos em DDS, identificando, de uma forma mais ampla, o que os participantes esperavam encontrar nesses estudos. Neste sentido, foram citados itens relacionados ao contexto da empresa (estrutura, localização geográfica, cultura), metodologia utilizada, instrumentos de coleta de dados e questão de pesquisa, contribuições do estudo para academia e indústria, perfil da população estudada (profissionais da indústria ou estudantes) e lições aprendidas.

Alguns trechos das entrevistas transcritas permitem ilustrar os resultados, como, por exemplo, este de um pesquisador do Brasil:

*“Eu espero encontrar a questão de pesquisa e as respostas para a mesma. É importante saber se alguém já respondeu a questão de pesquisa em outros estudos. [...] Em que contexto o trabalho foi feito, ou seja, uma descrição detalhada da empresa, que tipo de negócio, privada ou pública, distribuição geográfica, estrutura organizacional (filosofia, políticas, forma de comunicação, cultura). [...] Que tipo de ferramenta a empresa e equipes*

*usam. Hoje em dia os processos podem estar embutidos no ferramental. Papéis e responsabilidades da equipe da empresa. [...] Como a pesquisa foi feita. Como foram selecionados os respondentes da pesquisa. Qual o perfil dos respondentes. Eram profissionais da indústria ou estudantes. Qual era o instrumento de coleta de dados.”*

E o relato de outro pesquisador do Brasil reforçando a importância do contexto das empresas e lições aprendidas:

*“O contexto das empresas e os resultados alcançados no estudo para que possa ser avaliado e comparado com outro estudo. [...] Lições aprendidas. [...] Nos casos de avaliações de ferramentas ou protótipos de ferramentas, seria importante ter detalhes da ferramenta ou protótipo, saber como ela foi testada, resultados de sucesso e melhorias. Em termos de modelo, ou seja, mais abstrato, seria importante ter uma avaliação qualitativa deste modelo.”*

A seguir é transcrita a expectativa de um pesquisador da Suécia a respeito da contribuição do estudo:

*“Contexto claro do estudo, incluindo a descrição da metodologia. [...] Resultados claros, com as implicações do estudo para academia e indústria.”*

Alguns respondentes relataram expectativas mais específicas, como encontrar resultados de sucesso e melhorias para avaliações de ferramentas ou protótipos de ferramentas e no caso de estudos que abordam aspectos de comunicação em DDS, por exemplo, é importante descrever como o experimento foi executado, que tipo de tarefas os times executaram, como os times estavam distribuídos e que tipo de software foi utilizado.

A síntese dos principais itens que os respondentes da academia esperam encontrar em estudos empíricos em DDS é apresentada na Tabela 16.

Tabela 16 – Síntese das expectativas da academia.

<b>Expectativas da academia com relação aos estudos empíricos</b>
Contexto da empresa
Método de pesquisa e instrumento de coleta de dados
Caracterização da população estudada (profissionais da indústria ou estudantes)
Contribuições do estudo para academia e indústria
Lições aprendidas

#### 4.3.1.2.3 Desafios

Com relação aos desafios inerentes ao processo de busca por artigos empíricos em DDS, cinco de um total de sete respondentes da academia, relataram que a busca por estudos empíricos é uma atividade com complexidade no mínimo alta, conforme o gráfico da Figura 15.

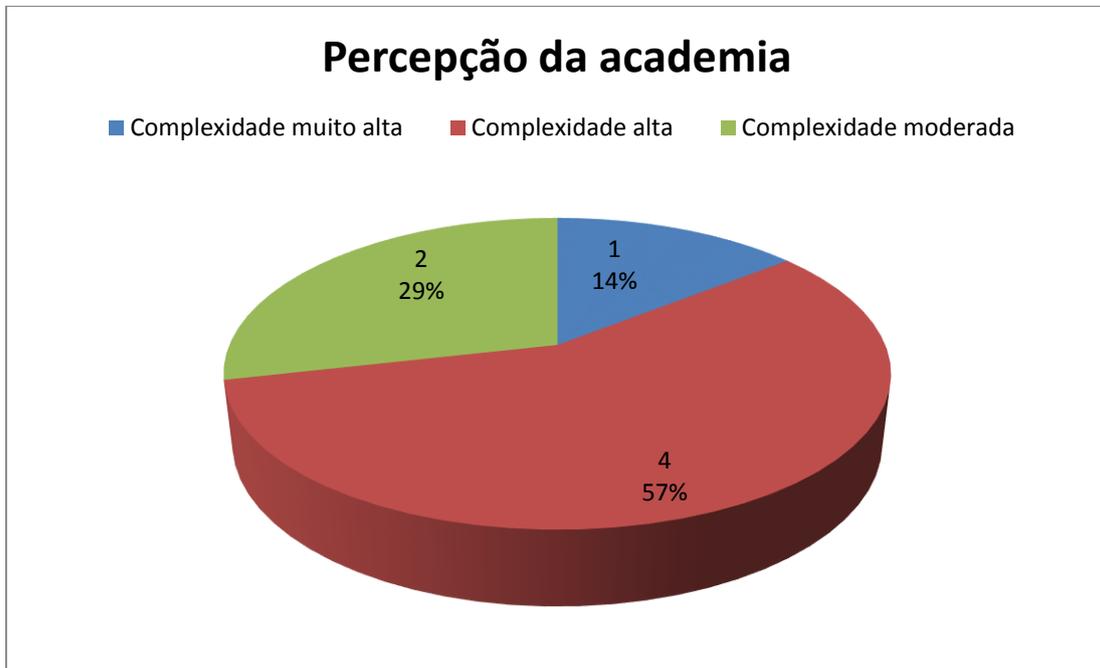


Figura 15 - Percepção da academia sobre a busca por artigos empíricos.

Os respondentes citaram diversos problemas relacionados com a descrição do contexto do estudo, terminologia da área, escolha das palavras-chaves, entre outros. A seguir são detalhados alguns desafios indicados pelos respondentes:

- Caracterização do estudo como empírico ou não empírico.
- Dificuldade em julgar a relevância e aplicabilidade da evidência empírica em outro contexto de estudo;
- Terminologia não padronizada;
- Dados empíricos, de uma perspectiva científica, são geralmente incompletos;
- Os *abstracts* dos artigos científicos não são estruturados, ou seja, muitas vezes não apresentam informações importantes como objetivos, como a pesquisa foi conduzida, resultados e a contribuição.
- Volume alto de artigos retornados que não fazem parte do objetivo da pesquisa;

- A maturidade da área influencia nos resultados;
- Mecanismos de pesquisa nas bases de dados.

Alguns trechos das entrevistas transcritas permitem ilustrar os resultados, como, por exemplo, a citação de um pesquisador do Brasil:

*"Não existem bons mecanismos de pesquisa nas bases de dados. O pesquisador muitas vezes prefere a busca manual devido às dificuldades encontradas em algumas bases de dados. [...] Utilizar String de busca também é complexo, pois os estudos não utilizam palavras chave no título e no abstract. Os abstracts não são estruturados e muitas vezes não apresentam informações importantes tais como: objetivo do artigo, como a pesquisa foi conduzida, resultado e a contribuição. Devido à falta de estrutura, o pesquisador é obrigado a ler todo o artigo. [...] Pesquisadores publicam artigos empíricos, sem mencionar que o estudo é empírico. Experimento também é empírico. É preciso ter uma classificação clara dos estudos empíricos e não empíricos. [...] Exemplo, o pesquisador avalia uma ferramenta através de um estudo controlado, mas não diz que foi empírico."*

A falta de informações de contexto dos estudos empíricos foi citada por alguns respondentes, conforme o relato deste pesquisador da Suécia:

*"Não é complexo procurar estudos empíricos, mas às vezes é difícil encontrar estudos relevantes ou julgar a relevância e aplicabilidade da evidência para o seu próprio contexto e isso ocorre principalmente porque muitas vezes o contexto do estudo não é bem descrito."*

Quanto à terminologia e a maturidade da área, outro pesquisador relatou:

*"Em DDS, existe a questão da terminologia não padronizada e isso dificulta principalmente quando o pesquisador é novo na área. Existem muitos termos para explicar a mesma coisa e isso torna a busca complexa. [...] Existem outros desafios relacionados com a maturidade da área, por exemplo, a área de teste dentro do contexto de DDS é carente de estudos empíricos que aborde aspectos técnicos na parte de qualidade. Muitas vezes o pesquisador não sabe mensurar se o volume de estudos retornados é pequeno porque os termos não estão corretos ou realmente não existem muitos estudos com esse tema."*

A seguir, apresenta-se na Tabela 17, a síntese dos principais desafios relatados com relação à busca por estudos empíricos em DDS:

Tabela 17 – Síntese dos desafios da academia.

Desafios da academia com relação à busca por estudos empíricos
Caracterização do estudo como empírico ou não empírico
Terminologia da área
Maturidade do campo de pesquisa
Resumo ( <i>abstract</i> ) não estruturado
Avaliar a relevância do estudo

#### 4.3.1.2.4 Contexto

O objetivo destas questões foi compreender que tipo de informação de contexto, em estudos empíricos em DDS, é importante para a academia e a indústria e dessa forma, identificar elementos de classificação que possam ajudar na proposta de uma taxonomia para área de DDS. A seguir, apresentam-se os resultados de cada artigo.

#### Artigo A:

As sentenças do artigo A, apresentadas na Figura 16, foram analisadas em ordem decrescente de relevância, ou seja, da mais relevante para menos relevante.

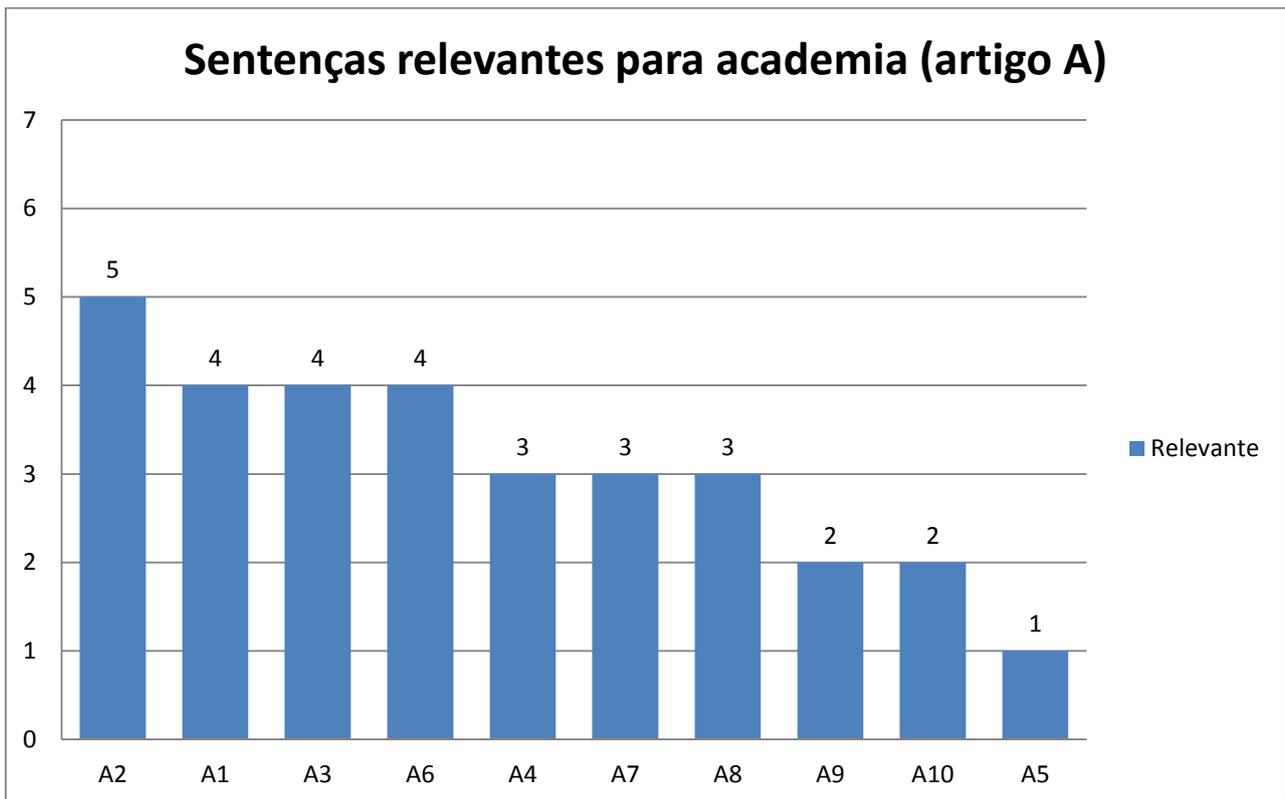


Figura 16 – Resultados da academia (artigo A).

A sentença A2 foi selecionada por cinco dos sete respondentes da academia, destacando-se como um elemento importante para compreender o contexto empírico do estudo, ou seja, os dados empíricos foram gerados a partir da indústria e não da

academia. As sentenças selecionadas por mais de 50% dos respondentes estão listadas a seguir:

- A2: *“The context for the research was the Latvian software development company LatSoftware, situated in Riga (Latvia in Europe)”*;
- A1: *“The objective of this article is to understand the importance of trust in GSD by describing the key factors that lead to a lack of trust, and the effect that a lack of trust has on a GSD team”*;
- A3: *“We collected data from supply teams that participated in four globally distributed software projects run by LatSoftware”*;
- A6: *“We used multiple data sources: qualitative interviews with project participants, results from postmortem meetings, and project problem reports”*.

As informações sobre o contexto da empresa, localização geográfica, método de coleta de dados e a caracterização dos projetos estudados estão entre as principais razões relatadas pelos respondentes. A seguir, descrevemos o relato de um pesquisador do Brasil:

*“Saber o país ajuda a entender os resultados, pois temos uma conotação atrelada à cultura. É importante saber se o país está influenciando no resultado da pesquisa. [...] É importante saber o número de projetos que foram estudados e suas características.”*

Com relação a informações de contexto que não ficaram claras, os respondentes citaram diversos pontos, tais como, o tamanho da equipe e como as pessoas estavam distribuídas nos projetos, que tipo de software foi desenvolvido, caracterização dos projetos analisados e a caracterização da população estudada. A seguir apresenta-se a opinião de um pesquisador da Itália:

*“Não está claro quem foram os entrevistados. Faltou informação do perfil dos entrevistados. [...] Informação sobre o projeto, tais como, tamanho e duração. Também faltaram detalhes sobre o que era o projeto, ou seja, contexto do projeto. [...] Informações sobre o time. O mesmo time trabalhou nos quatro projetos? Como eles estavam distribuídos? Por quanto tempo o time trabalha junto? Já se encontraram face to face? Que tipo de ferramenta eles usaram para as reuniões?”*

### **Artigo B:**

Da mesma forma, as sentenças do artigo B são apresentadas na Figura 17:

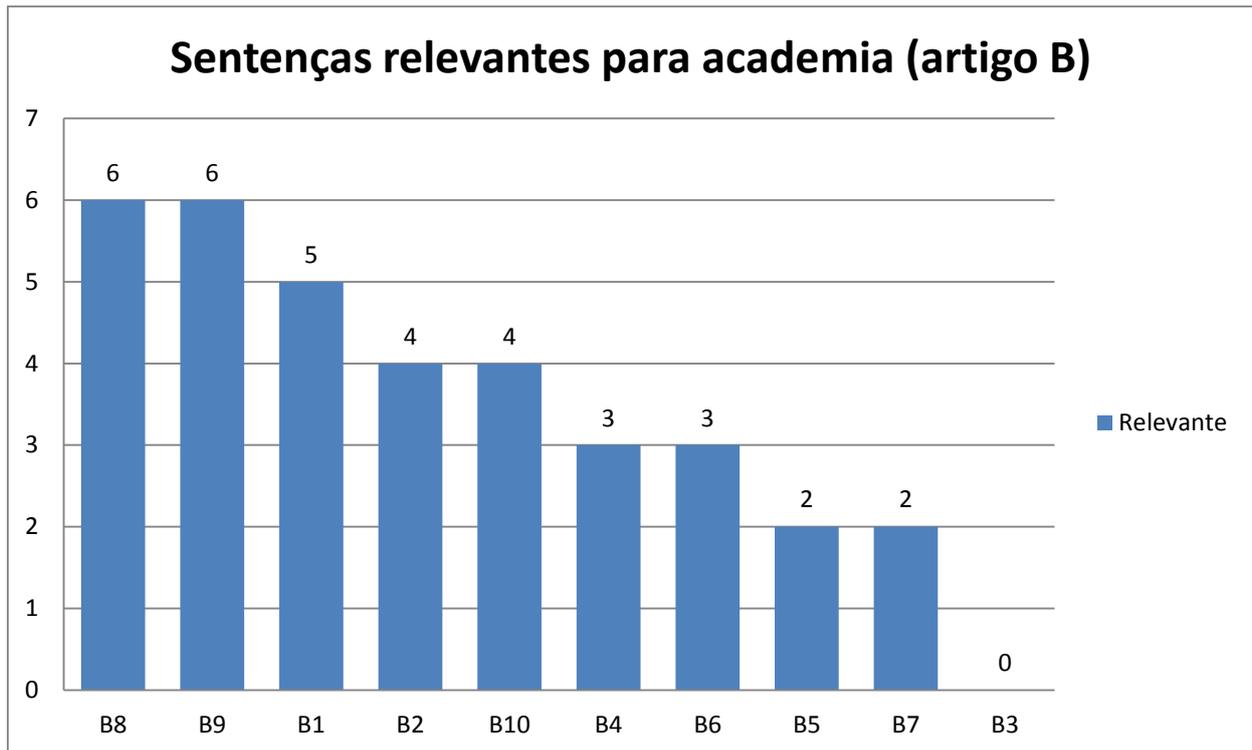


Figura 17 – Resultados da academia (artigo B).

A seguir, listamos as sentenças selecionadas por mais de 50% dos respondentes:

- *B8: “The team size varies from two to nine persons. Some teams are distributed between Norway and Malaysia. Whether a team is distributed or not can vary between iterations”;*
- *B9: “The time difference between the two sites is seven hours during the winter and six hours during the summer”;*
- *B1: “This paper reports a case study on agile practices in a 40-person development organization distributed between Norway and Malaysia”;*
- *B2: “Based on seven interviews in the development organization, we describe how scrum practices were successfully applied”;*
- *B10: “The main contribution of this paper is a detailed description on how scrum practices were successfully applied to a mid-sized distributed software product development program”.*

Os entrevistados destacaram que informações sobre o time de projeto, tais como, tamanho do time, facilitam a comparação com outros estudos. A distribuição dos times e as informações sobre distancia temporal ajudam a entender o contexto do estudo. O método de pesquisa e o instrumento de coleta de dados também foram destacados como

uma informação de contexto relevante para academia. O trecho transcrito a seguir, relata a opinião de um pesquisador da Itália:

*“A primeira sentença é útil, pois descreve o método utilizado, número de pessoas envolvidas e a distribuição do projeto de forma clara. O tamanho do time e a distribuição do mesmo são muito importantes, pois ajudam a comparar os resultados. A informação de quando o time está distribuído ou não também é relevante para entender o contexto do estudo. [...] Informação sobre diferença de tempo e distância temporal também são importantes.”*

Os respondentes relataram algumas informações de contexto que não ficaram claras, tais como, tamanho da organização, distribuição dos times e quais práticas de *Scrum* foram analisadas. O relato de um pesquisador da Suécia ilustra os resultados analisados:

*“As informações de contexto da primeira e da última sentença não estão claras em relação ao tamanho do produto. Como uma organização de desenvolvedores com 40 pessoas podem desenvolver um produto grande e ao mesmo tempo ser classificada como uma organização de médio porte? [...] Como as 40 pessoas estão distribuídas em duas localizações? Quanto tempo à organização existe? Porque ela foi iniciada? Pareceu que os times não são estáveis, por quê? O que eles desenvolvem? [...] Os dois sites usam agile desde o início ou ambos transicionaram de um modelo tradicional de desenvolvimento?”*

A seguir, apresenta-se na Tabela 18, a síntese das informações de contexto mais relevantes para academia. Como resultado desta análise, consideraram-se, também, os relatos sobre as informações que não estavam claras nos estudos.

Tabela 18 – Síntese das informações de contexto (academia).

Informações de contexto relevantes para academia
Contexto empírico do estudo (indústria ou laboratório)
Caracterização da organização (localização geográfica, tamanho)
Caracterização dos projetos estudados (distribuição dos times e distância temporal)
Caracterização da população estudada
Método de pesquisa
Método de coleta de dados (entrevista, observação, etc.)

### 4.3.2 Resultados da indústria

A seguir serão apresentadas as caracterizações dos respondentes e os elementos de análise, assim como os resultados obtidos dos profissionais da indústria.

#### 4.3.2.1 Caracterização dos respondentes da indústria

Na indústria, oito profissionais foram entrevistados, sendo estes executivos, gerentes, líderes técnicos e pesquisadores que atuam na indústria. Deste total, cinco entrevistas foram presenciais e três participantes enviaram as respostas do questionário por e-mail, devido à falta de disponibilidade para realização da entrevista. As entrevistas tiveram uma duração média de 32 minutos (entre um mínimo de 25 minutos e um máximo de 40 minutos).

Com relação ao tempo de experiência na indústria, o grupo possui uma grande maturidade em relação ao DDS, sendo que metade dos respondentes possuem mais de 10 anos de experiência, três profissionais com experiência entre 6 a 10 anos e apenas um entre 2 e 4 anos (Figura 18).

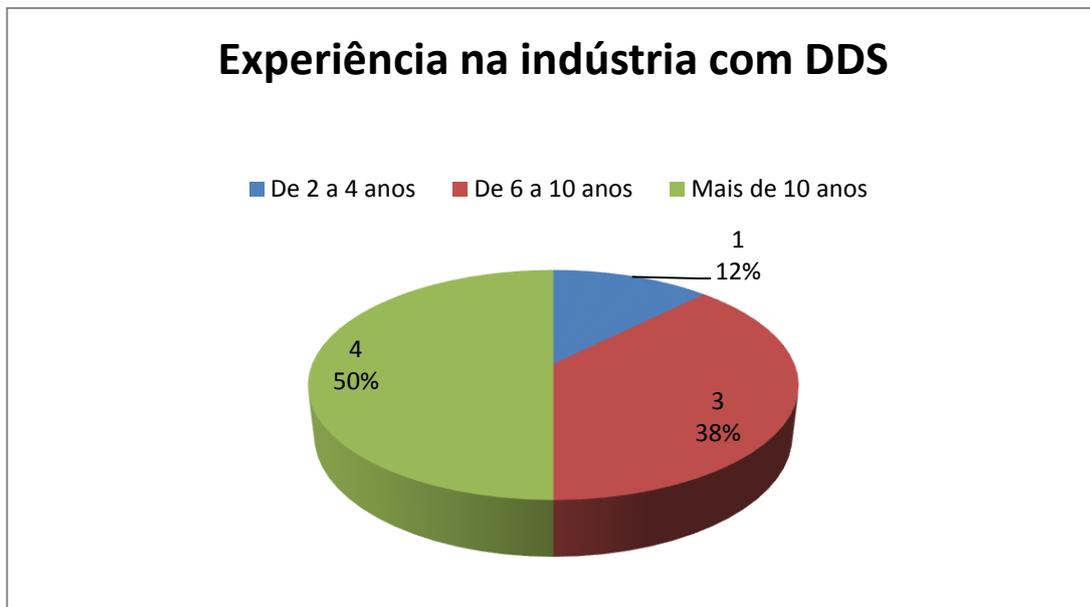


Figura 18 – Tempo de experiência dos profissionais da indústria com DDS.

Quanto à experiência com pesquisa em DDS, 50% dos respondentes possuem de dois a quatro anos, sendo que três profissionais possuem mais de 10 anos. Dos oito respondentes, cinco possuem o título de mestre e um o título de doutor, confirmando novamente o alto nível de qualificação do grupo. Quanto à distribuição geográfica, a

Figura 19, apresentada a seguir, mostra que cinco dos oito respondentes da indústria trabalham no Brasil e os demais estão localizados na Índia e nos Estados Unidos.

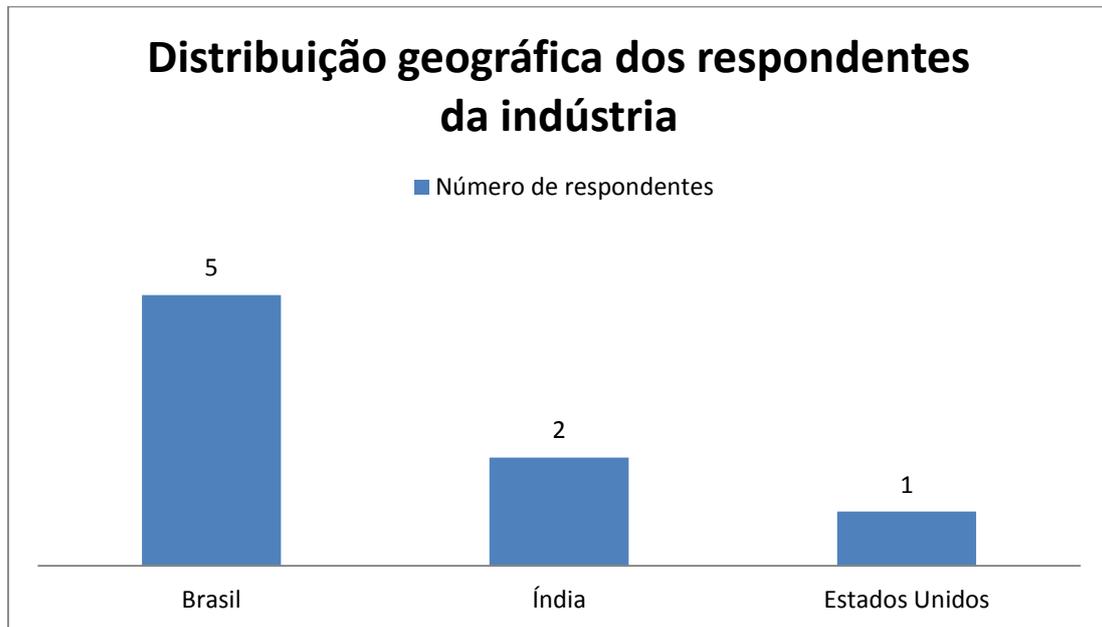


Figura 19 - Distribuição geográfica dos respondentes da indústria.

#### 4.3.2.2 Elementos de análise da indústria

A seguir, apresentam-se os elementos analisados e as categorias obtidas, buscando direcionar para um conjunto de lições relevantes visando contemplar o objetivo principal deste estudo, relacionado com a busca de uma taxonomia para sistematização das evidências empíricas em DDS.

##### 4.3.2.2.1 Motivações

Essa questão buscava compreender, de uma forma mais ampla, quais seriam as finalidades ou motivações nas quais os respondentes da indústria buscariam estudos empíricos na literatura relacionados com a área de pesquisa em DDS.

Os entrevistados citaram diversas motivações pelas quais buscariam artigos empíricos na literatura, tais como, melhores práticas, benchmarking e soluções para problemas ou dificuldades inerentes ao DDS. Alguns trechos das entrevistas transcritas ajudam a ilustrar os resultados, como, por exemplo, este de um profissional da Índia com mais de 10 anos de experiência na área:

*“Para melhorar as nossas operações descobrindo melhores práticas de outras empresas e equipes que utilizam DDS e que conseguiram superar os desafios enfrentados em suas organizações.”*

E o relato de um profissional do Brasil que atua em uma multinacional e possui mais de 10 anos de experiência em DDS:

*“Encontrar boas práticas em DDS, principalmente em relação a times distribuídos e como eles podem trabalhar juntos de forma mais efetiva. [...] Encontrar estudos de casos com equipes distribuídas utilizando métodos ágeis.”*

A seguir, apresenta-se na Tabela 19, a síntese das principais motivações relatadas pela indústria em relação à busca por estudos empíricos em DDS.

Tabela 19 – Síntese das motivações da indústria para busca de estudos empíricos.

Motivações da indústria para busca de estudos empíricos em DDS
Melhores práticas
Benchmarking
Soluções para problemas inerentes ao DDS

#### 4.3.2.2.2 Expectativas

Buscou-se explorar com essa questão, o que os participantes esperavam encontrar nesses estudos. Neste sentido, os respondentes relataram que gostariam de encontrar resultados de sucesso ou falha dos estudos, informações sobre custo e benefício de uma determinada prática ou solução, relatos de desafios em DDS e como eles foram resolvidos, lições aprendidas e metodologia aplicada no estudo. A seguir é transcrita a expectativa de um profissional do Brasil que trabalha em uma multinacional americana:

*“Informação visual simples, como gráficos, que sumarizem o problema e a solução. [...] Encontrar estudos empíricos com o mesmo cenário da minha empresa. É importante categorizar a empresa para poder comparar problemas e soluções. [...] Informações sobre custo e benefício de uma determinada prática ou solução.”*

Alguns profissionais relataram expectativas com relação a soluções para desafios em DDS, como citou este executivo que trabalha em uma multinacional na Índia:

*“Ideias sobre os desafios enfrentados por outras equipes DDS. Explicações sobre por que esses desafios surgem em DDS. Quais medidas foram tomadas para enfrentar esses desafios. Avaliações objetivas dos principais benefícios alcançados.”*

E o relato de um profissional da indústria que trabalha em uma multinacional nos Estados Unidos sobre resultados de sucesso ou falha dos estudos:

*“Informação sobre o projeto estudado, como local, tamanho, objetivo. [...] Quanto aos resultados do estudo, sucesso ou falha do projeto. [...] Informações sobre a cultura da empresa. [...] Dados sobre a estrutura de gerencia da empresa. [...] Evidências sobre o que funcionou e o que não funcionou no estudo.”*

A Tabela 20 apresenta a síntese dos principais itens que os respondentes da indústria esperam encontrar em estudos empíricos em DDS.

Tabela 20 – Síntese das expetativas da indústria.

Expectativas da indústria com relação aos estudos empíricos
Contexto do projeto estudado
Método de pesquisa e instrumento de coleta de dados
Resultados de sucesso ou falha
Lições aprendidas
Custo e benefício de uma determinada prática ou solução
Relatos de desafios em DDS e como eles foram resolvidos

#### 4.3.2.2.3 Desafios

A primeira questão buscou mapear na, indústria e na academia, como os profissionais que trabalham e pesquisam em DDS percebem a complexidade do processo de busca por artigos empíricos nesta área. Metade dos respondentes da indústria relatou que a busca por estudos empíricos é uma atividade, no mínimo, com complexidade alta, enquanto que a outra metade considerou moderada (Figura 20).

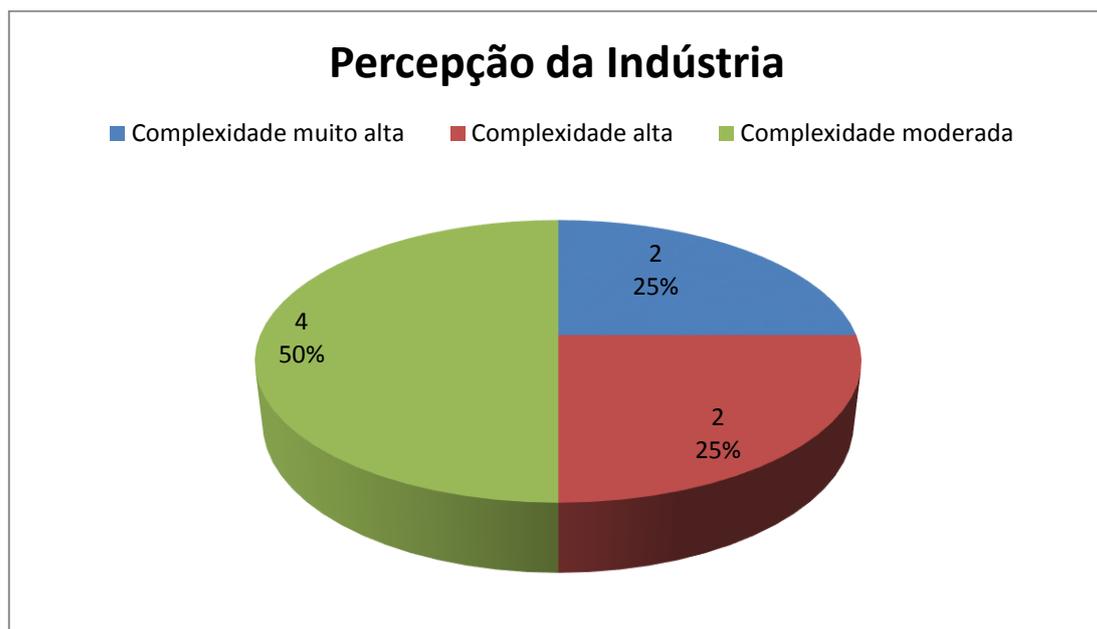


Figura 20 - Percepção da indústria sobre a busca por artigos empíricos.

Com relação aos desafios, os respondentes citaram diversos problemas relacionados com a terminologia da área, a dificuldade em avaliar a relevância dos estudos e com a qualidade do estudo publicado. A terminologia foi citada por alguns respondentes, conforme o relato do pesquisador do Brasil:

*“A complexidade é alta, pois o volume de estudos que não fazem parte da pesquisa gera muito ruído no processo de busca. [...] A terminologia da área pode dificultar o acesso aos artigos de interesse. [...] Por exemplo, o termo “Follow the Sun” retorna muito ruído nas bases de pesquisa devido à palavra sun.”*

Este gerente também citou a questão da terminologia da área:

*“A terminologia da área aumenta a complexidade desta tarefa, pois ela muda com o passar dos anos, devido à evolução natural da área. Por exemplo, a terminologia utilizada na década de 30 até a década de 50 para abordar o tema de terceirização entre empresas é diferente da terminologia utilizada na década de 50 até a década de 80, que por sua vez, é diferente da terminologia dos anos 90. [...] Este processo se torna mais complexo quando não conhecemos quem são os principais autores da área.”*

E um pesquisador da Índia que atua em uma multinacional e relatou a questão da qualidade dos estudos disponíveis na literatura:

*“A complexidade depende de vários fatores: (a) se o tema de pesquisa já foi pesquisado anteriormente; (b) qualidade da pesquisa feita até agora, especialmente a pesquisa empírica de empresas similares; (c) quão bem o tópico tem sido coberto na literatura; e (d) quão acessíveis são as fontes bibliográficas, pois às vezes enfrentamos problemas para obter financiamento para acesso a fontes pagas de literatura. [...] Já enfrentei problemas em torno de (a), (b) e (c). Especialmente para (c), pois comparar o trabalho com outras empresas é difícil. [...] Muitas vezes o contexto do estudo não é relatado na literatura acadêmica e muito provavelmente para proteger a confidencialidade da empresa.”*

E o relato de um líder de projeto que comenta sobre a dificuldade em julgar a relevância do estudo:

*“Hoje temos um excesso de material disponibilizado na web e não sabemos se encontramos os artigos mais importantes. [...] Eu entendo que a*

*complexidade está em selecionar os artigos que agregam valor, devido ao volume grande de estudos retornados.”*

A seguir, apresenta-se na Tabela 21, a síntese dos principais desafios relatados pela indústria com relação à busca por estudos empíricos em DDS:

Tabela 21 – Síntese dos desafios da indústria.

Desafios da indústria com relação à busca por estudos empíricos
Terminologia da área
Avaliar a relevância do estudo
Qualidade do estudo

#### 4.3.2.2.4 Contexto

Conforme descrito nas seções anteriores, estas questões buscavam identificar que tipo de informação de contexto dos estudos em DDS é relevante para profissionais e pesquisadores desta área. A seguir, apresentam-se os resultados de cada artigo.

#### **Artigo A:**

As sentenças do artigo A, apresentadas na Figura 21, foram analisadas em ordem decrescente de relevância, ou seja, da mais relevante para menos relevante.

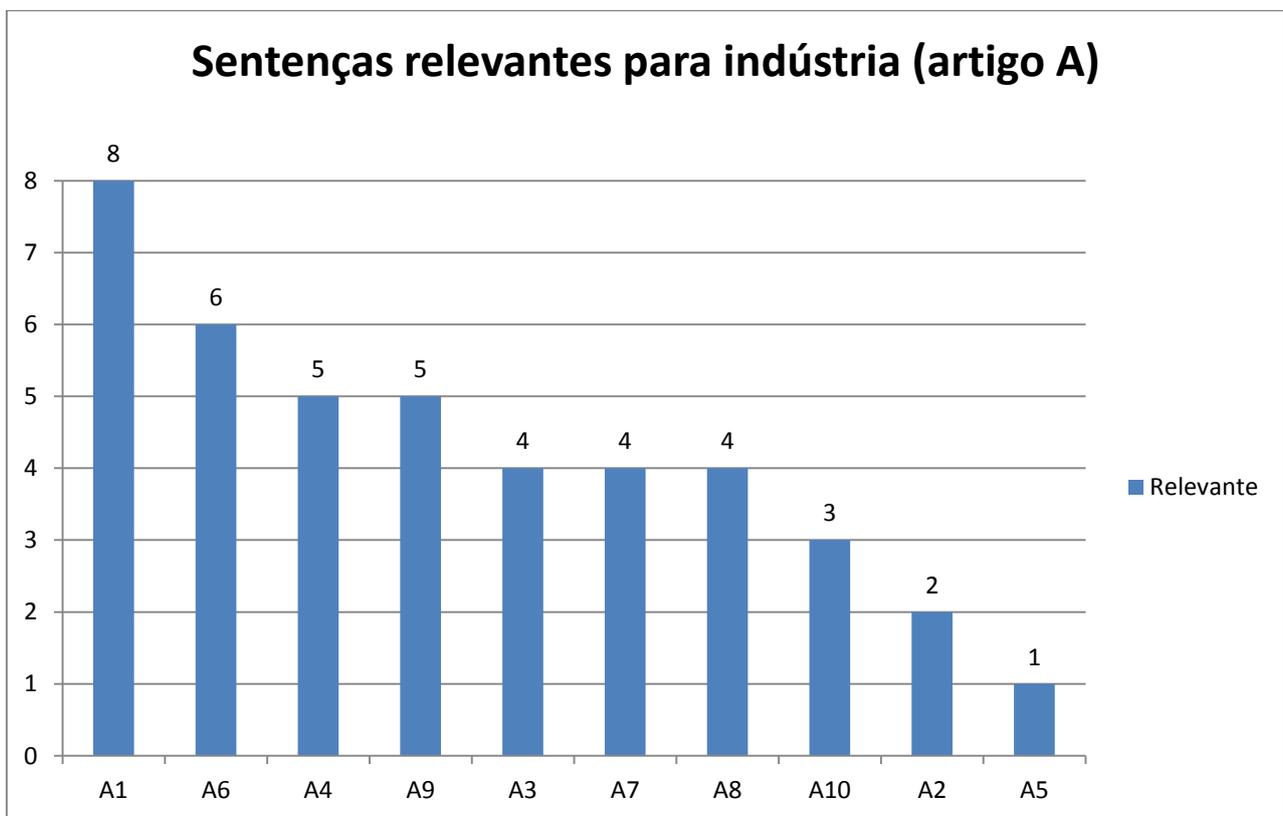


Figura 21 – Resultados da indústria (artigo A).

A sentença A1 foi selecionada por todos os respondentes da indústria, como sendo a mais relevante. A sentença A4 traz informações do contexto empírico do estudo, descrevendo que as evidências foram fundamentadas na indústria, através da análise de projetos globais de desenvolvimento de software. A seguir, listamos as sentenças selecionadas por mais de 50% dos respondentes:

- A1: *“The objective of this article is to understand the importance of trust in GSD by describing the key factors that lead to a lack of trust, and the effect that a lack of trust has on a GSD team”;*
- A6: *“We used multiple data sources: qualitative interviews with project participants, results from postmortem meetings, and project problem reports”;*
- A4: *“We choose four global software development projects that all reported problems with trust”;*
- A9: *“For the analysis, we relied mainly on qualitative interviews, because these provide a rich picture of the reasons for, and effects of, a lack of trust”.*

As informações sobre o objetivo do estudo, caracterização dos projetos estudados e instrumentos de coleta de dados foram as principais razões relatadas pelos respondentes. A seguir, descrevemos o relato de um profissional do Brasil:

*“As sentenças selecionadas descrevem o objetivo e a metodologia utilizada. Contexto e número de projetos estudados. [...] A metodologia aplicada para coleta de dados é importante para evidenciar que foram utilizados dados empíricos. [...] É importante descrever o método científico, como entrevistas qualitativas, que trouxeram resultados e atingiram o objetivo do estudo.”*

E o relato de outro profissional do Brasil reforçando a importância do objetivo do estudo e da metodologia aplicada:

*“A primeira sentença descreve o contexto do problema que o artigo quer investigar e isso ajuda a determinar rapidamente se o artigo é do meu interesse. [...] É importante saber que foi estudado vários casos com o mesmo problema. [...] A metodologia usada para coleta de dados mostra que o artigo apresentou evidências empíricas para entender o fenômeno de confiança nos times.”*

Com relação a informações de contexto que não ficaram claras, os entrevistados citaram diversos pontos, tais como, o tamanho da equipe e como as pessoas estavam distribuídas nos projetos, que tipo de software foi desenvolvido, caracterização dos

projetos analisados e a caracterização da população estudada. A seguir apresenta-se a opinião de um pesquisador da Itália:

*“Não está claro quem foram os entrevistados. Faltou informação do perfil dos entrevistados. [...] Informação sobre o projeto, tais como, tamanho e duração. Também faltaram detalhes sobre o que era o projeto, ou seja, contexto do projeto. [...] Informações sobre o time. O mesmo time trabalhou nos quatro projetos? Como eles estavam distribuídos? Por quanto tempo o time trabalha junto? Já se encontraram face to face? Que tipo de ferramenta eles usaram para as reuniões?”*

### Artigo B:

Da mesma forma, as sentenças do artigo B são apresentadas na Figura 22:

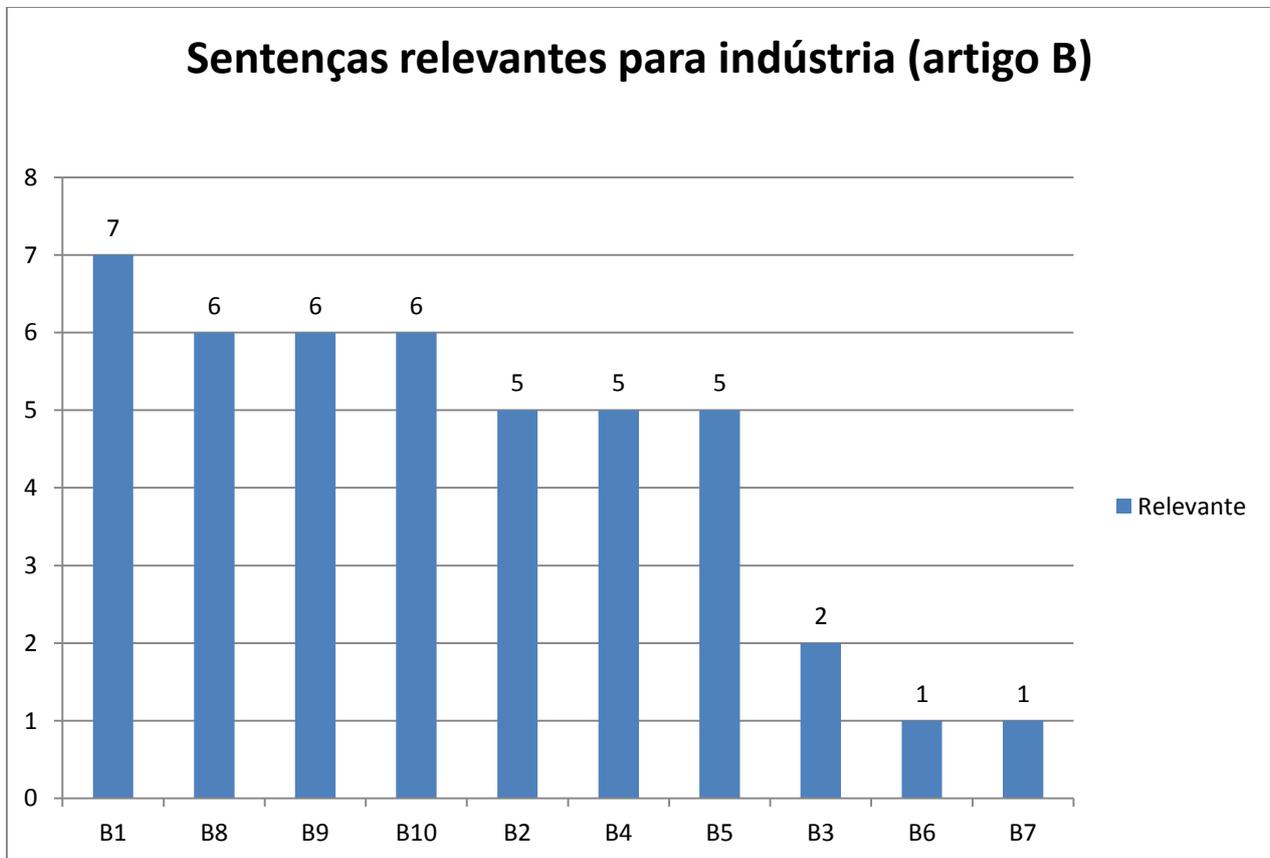


Figura 22 – Resultados da indústria (artigo B).

A sentença B1, que descreve o método principal de pesquisa e o contexto empírico do estudo, destacou-se como sendo a sentença mais relevante do estudo B, ou seja, o método de pesquisa foi um estudo de caso e as evidências empíricas foram fundamentadas na indústria. A seguir, listamos as quatro sentenças mais relevantes para indústria:

- B1: *“This paper reports a case study on agile practices in a 40-person development organization distributed between Norway and Malaysia”;*
- B8: *“The team size varies from two to nine persons. Some teams are distributed between Norway and Malaysia. Whether a team is distributed or not can vary between iterations”;*
- B9: *“The time difference between the two sites is seven hours during the winter and six hours during the summer”;*
- B10: *“The main contribution of this paper is a detailed description on how scrum practices were successfully applied to a mid-sized distributed software product development program”.*

Com relação às sentenças mais relevantes, os profissionais da indústria relataram que o tamanho e a distribuição do time de projeto facilitam a comparação do estudo com outros cenários. A contribuição do estudo também foi outra informação relevante para indústria. O trecho transcrito a seguir, relata a opinião de um profissional do Brasil:

*“O tamanho da organização, tamanho do time e a distribuição do time de projeto são informações de contexto importantes para indústria, pois facilitam a comparação com outros cenários. [...] assim como o número de estudos de caso realizados.”*

Assim como, o relato de outro profissional do Brasil sobre a distribuição do time de projeto:

*“A primeira sentença traz o tamanho da equipe e a distribuição nos países, abordando de forma resumida o contexto do estudo. [...] Distribuição do time de projeto em dois países. Essa informação de contexto é muito importante para um estudo em DDS, pois existem diversos fatores como cultura, distancia geográfica, língua, entre outros, que podem influenciar no resultado do estudo. [...] Tamanho da equipe é uma informação de contexto muito importante para projetos ágeis, assim como, diferenças de fuso horário. [...] Essas informações também são importantes para fins de comparação com outros estudos.”*

Os respondentes relataram algumas informações de contexto que não ficaram claras, tais como, tamanho da organização, maturidade do time de projeto com relação a métodos ágeis e informações sobre os resultados do estudo, incluindo detalhes sobre quais práticas do *Scrum* tiveram sucesso e quais não funcionaram. O relato de um profissional do Brasil ilustra os resultados analisados:

*“Não ficou claro que significa ‘um grande programa distribuído’. Poder ser um programa com 500 pessoas ou pode ser um programa com 50. [...] O que significa uma empresa de médio porte? A primeira sentença fala sobre um time com 40 pessoas e depois diz que o programa é médio.”*

A seguir, apresenta-se na Tabela 22, a síntese das informações de contexto mais relevantes para indústria. Como resultado desta análise, consideraram-se, também, os relatos sobre as informações que não estavam claras nos estudos.

Tabela 22 – Síntese das informações de contexto (indústria).

Informações de contexto relevantes para indústria
Contexto empírico do estudo (indústria ou laboratório)
Caracterização da organização (tamanho e localização)
Caracterização dos projetos estudados (distribuição dos times e distância temporal)
Método de pesquisa
Método de coleta de dados (entrevista, observação, etc.)

#### 4.4 CONSOLIDAÇÃO DOS RESULTADOS DA SURVEY

A seguir, apresenta-se na Tabela 23, a consolidação dos resultados da academia e da indústria, conforme as quatro categorias analisadas na *survey*: motivações, expectativas, desafios e contexto dos estudos empíricos em DDS.

Na primeira categoria, motivação, buscou-se compreender para qual finalidade os respondentes buscariam artigos empíricos na literatura relacionados com a área de pesquisa em DDS. Essa questão enriqueceu esta pesquisa, na medida em que explorou novas ideias e temas para pesquisas futuras.

Com relação à segunda categoria, foram explorados quais eram as expectativas dos respondentes com relação aos estudos empíricos em DDS, identificando, de uma forma mais ampla, o que os participantes esperavam encontrar nesses estudos. Os resultados mostraram que a descrição do contexto das empresas e dos projetos estudados, a metodologia do estudo e as lições aprendidas são aspectos importantes para os profissionais que trabalham e pesquisam em DDS.

Com base nos resultados deste estudo, ambos, pesquisadores e profissionais da indústria, relataram a terminologia da área e a dificuldade em avaliar a relevância de um estudo, como desafios inerentes ao processo de busca por estudos empíricos na área. A falta de caracterização dos estudos como empírico ou não empírico também foi citada como uma dificuldade relacionada a este processo.

Tabela 23 – Consolidação dos resultados da *survey*

<b>Categoria</b>	<b>Resultado</b>	<b>Fonte</b>
<b>Motivações</b>	Fundamentação teórica	Academia
	Responder as perguntas da indústria	Academia
	Soluções e melhores práticas	Academia e Indústria
	Conhecer o estado da arte da área	Academia
	Benchmarking	Indústria
<b>Expectativas</b>	Contexto das empresas e dos projetos estudados	Academia e Indústria
	Método de pesquisa e instrumento de coleta de dados	Academia e Indústria
	Caracterização da população estudada	Academia
	Contribuições do estudo	Academia
	Lições aprendidas	Academia e Indústria
	Resultados de sucesso ou falha	Indústria
	Custo e benefício de uma determinada prática ou solução	Indústria
	Relatos de desafios em DDS e como eles foram resolvidos	Indústria
<b>Desafios</b>	Caracterização do estudo (empírico ou não empírico)	Academia
	Terminologia da área	Academia e Indústria
	Maturidade do campo de pesquisa	Academia
	Resumo ( <i>abstract</i> ) não estruturado	Academia
	Avaliar a relevância do estudo	Academia e Indústria
	Qualidade do estudo	Indústria
<b>Contexto</b>	Contexto empírico do estudo (indústria ou laboratório)	Academia e Indústria
	Caracterização da organização	Academia e Indústria
	Caracterização dos projetos estudados	Academia e Indústria
	Caracterização da população estudada	Academia
	Método de pesquisa e método de coleta de dados	Academia e Indústria

Por fim, na categoria contexto, explorou-se que tipo de informação de contexto, dos estudos empíricos em DDS, é relevante para profissionais e pesquisadores desta área. O contexto empírico do estudo (academia ou indústria), a categorização das organizações e dos projetos estudados e a metodologia de pesquisa (método principal e instrumentos de coleta de dados) foram mencionados, pelos respondentes, como sendo elementos importantes para descrição do contexto de um estudo empírico.

## 4.5 LIÇÕES PARA O ESTUDO

Os resultados da academia e da indústria ilustraram diversos aspectos referentes às expectativas, desafios e informações de contexto que são importantes para profissionais e pesquisadores em DDS. As lições aprendidas apresentadas a seguir estão baseadas no confronto entre a teoria e as descobertas empíricas destes estudos, e servirão como base de sustentação para a taxonomia proposta nesta pesquisa.

### **Lição 1: A indústria procura soluções e melhoras práticas em DDS para comparar os resultados com seu próprio contexto**

Soluções para os problemas inerentes ao DDS e melhores práticas, estão entre os principais motivos pelos quais profissionais da indústria buscam estudos empíricos na literatura, com a finalidade de comparar os resultados com seu próprio contexto organizacional e de projetos distribuídos. Por outro lado, a academia está focada em encontrar respostas para as perguntas dos parceiros da indústria, entre outros motivos, mostrando indícios claros de que a academia e a indústria estão alinhadas com relação ao campo de estudo em DDS.

### **Lição 2: Existe a necessidade de caracterizar se o estudo é de natureza empírica ou não**

Os resultados deste estudo mostraram que muitas vezes os pesquisadores publicam artigos de natureza empírica, sem mencionar que o estudo é empírico. Esse tipo de informação facilitaria o processo de busca por evidências empíricas na literatura, uma vez que pesquisadores e profissionais da indústria poderiam filtrar os estudos pelas palavras chave.

### **Lição 3: É importante identificar o contexto empírico no qual as evidências do estudo estão fundamentadas**

Através dos artigos empíricos analisados na *survey*, constatou-se que as sentenças descrevendo o contexto empírico de cada estudo estavam entre as informações de contexto mais relevantes para pesquisadores e profissionais da indústria. Dessa forma, é importante relatar se os dados empíricos estão baseados na indústria ou em experimentos de laboratório, fornecendo informações claras do contexto da pesquisa para que os profissionais da indústria e pesquisadores possam interpretar melhor os resultados do estudo.

#### **Lição 4: O método de pesquisa e os métodos de coleta de dados empíricos são informações de contexto importantes para indústria e academia**

Entre as informações de contexto mais relevantes para os respondentes deste estudo estão os métodos de pesquisa e o método de coleta de dados, tais como, entrevistas, resultados das reuniões *post-mortem*, relatórios de problemas de projeto, entre outros. Os respondentes relataram que a descrição do método aplicado para coleta de dados é importante para evidenciar que foram utilizados dados empíricos no estudo.

#### **Lição 5: A caracterização da população estudada ajuda na compreensão e na interpretação dos resultados do estudo**

A caracterização da população estudada foi apontada como um fator importante para compreensão dos resultados, uma vez que envolvem aspectos culturais, maturidade da equipe, tamanho e nível de distribuição do time de projeto, entre outros. Os respondentes relataram, por exemplo, que é necessário identificar se a população estudada é composta por estudantes ou profissionais da indústria.

#### **Lição 6: A caracterização das organizações e dos projetos estudados ajuda a avaliar a aplicabilidade do estudo em outros contextos**

A caracterização das organizações e os projetos estudados foram identificados, pela indústria e academia, como informações essenciais para julgar a aplicabilidade do estudo em outros contextos e compreender os resultados empíricos do estudo. Os respondentes relataram, por exemplo, que a localização geográfica pode influenciar nos resultados da pesquisa. Outras informações como a distribuição do time de projeto, tamanho e distância temporal também apareceram nos resultados como informações de contexto importantes para o entendimento do estudo.

#### **Lição 7: A padronização da terminologia da área é um fator importante para facilitar a busca por evidências empíricas em DDS**

Os resultados mostraram que a terminologia é um fator limitante para a busca de estudos empíricos, principalmente quando o pesquisador é novo na área. Os respondentes relatam que existem muitos termos para explicar o mesmo fenômeno e isso torna a busca por estudos em DDS uma tarefa complexa. Os participantes do estudo observaram que a terminologia utilizada na década de 30 até a década de 50, para abordar o tema de terceirização, é diferente da terminologia utilizada na década de 50 até a década de 80, que por sua vez, é diferente da terminologia dos anos 90.

**Lição 8: A avaliação sobre sucesso ou falha, dos resultados obtidos no estudo, é uma informação importante para os profissionais da indústria**

A descrição sobre sucesso ou falha dos resultados do artigo foi constatada na *survey* como uma expectativa dos profissionais da indústria, quando estes procuram estudos empíricos em DDS na literatura. Este tipo de informação pode auxiliar na classificação dos estudos empíricos e conseqüentemente, facilitar o processo de busca por evidências empíricas de sucesso ou falha dentro do contexto de DDS.

## 5 TAXONOMIA

Uma das dificuldades que pesquisadores e profissionais da indústria enfrentam está relacionada com o processo de deduzir e coletar informações sobre o contexto empírico dos estudos analisados, conforme visto nos resultados apresentados no Capítulo 4. Isto justifica a necessidade de descrições mais completas do contexto em que os estudos são realizados. A partir dos conhecimentos adquiridos nos estudos de base teórica, indicados no Capítulo 2, e as lições aprendidas da *survey* com especialistas, descrita no Capítulo 4, esta pesquisa propõe uma taxonomia para sistematização das evidências empíricas em Desenvolvimento Distribuído de Software.

Uma taxonomia é uma estrutura categorizada e a classificação é a ação de atribuição de entidades às categorias definidas dentro da taxonomia [SMI12]. É o agrupamento de itens semelhantes, tomando por base critérios estabelecidos [PRIE02]. Segundo Smite et al. [SMI12], o principal objetivo de uma taxonomia é ajudar os pesquisadores a aplicar suas pesquisas sistematicamente e facilitar a comparação e aplicação adequada da pesquisa dentro da área de DDS. A taxonomia também pode ser usada na aplicação prática dos resultados da investigação. Por exemplo, resultados de pesquisas para uma determinada dimensão de DDS podem ser categorizadas dentro da taxonomia, facilitando, dessa forma, a sistematização do conhecimento. Forward e Lethbridge [FOR08] complementam que uma taxonomia pode ser usada para ajudar a categorizar estudos empíricos e outras evidências obtidas por pesquisadores de DDS, de modo que a aplicabilidade da evidência pode ser mais sistematicamente compreendida.

Dessa forma, a taxonomia proposta visa auxiliar pesquisadores e profissionais da indústria na compreensão de como as evidências empíricas devem ser relatadas. A taxonomia pode ser usada como um referencial teórico-prático, buscando melhorar a publicação de estudos empíricos e relatos de experiência, assim como, a busca por estudos empíricos em DDS.

A seção 5.1 apresenta uma descrição sobre as categorias e os elementos de classificação da taxonomia e na seção 5.2 apresenta-se o processo de avaliação da mesma. Por fim, na seção 5.3 descrevem-se os resultados da avaliação da taxonomia.

## 5.1 CATEGORIAS E ELEMENTOS DA TAXONOMIA PROPOSTA

A seguir serão apresentadas as categorias e os elementos identificados nesta pesquisa, através de uma revisão aprofundada da base teórico e dos resultados obtidos na *survey* realizado com pesquisadores e profissionais da indústria.

A taxonomia proposta, ilustrada na Figura 23, é formada por onze categorias. Cada categoria contém elementos que auxiliam na classificação das evidências empíricas geradas a partir de estudos científicos em DDS. A taxonomia deve ser lida da esquerda para direita, começando pelo primeiro elemento que identifica os estudos que apresentam evidências empíricas, ilustrado na categoria “DDS”. Conforme indicado na legenda da Figura 23, os elementos de cada categoria são representados por um retângulo e possuem uma ligação direta com todos os elementos da próxima categoria. Os elementos que contém um sinal de “+” podem se conectar com todos os elementos da próxima categoria. As categorias que definem a localização da terceirização (*sourcing*) de um serviço ou produto não se aplicam quando o sujeito do estudo são estudantes, como é o caso do elemento “Estudantes” da categoria “Sujeito do estudo”. A seguir, descrevemos as categorias e os elementos de classificação da taxonomia:

- A primeira categoria representa o ponto inicial da taxonomia e contém apenas um elemento: evidências empíricas. Este elemento reforça a necessidade de caracterização do estudo como empírico, conforme evidenciado nos resultados da *survey*. Buscou-se como parte desta pesquisa analisar apenas estudos com evidências empíricas em DDS e, portanto, estudos de natureza não empírica foram excluídos dos objetivos desta pesquisa.
- A segunda categoria da taxonomia identifica se as evidências empíricas estão fundamentadas em dados da indústria ou em dados gerados a partir de um experimento de laboratório. A necessidade de especificar o contexto empírico foi evidenciada na *survey*, através da análise dos artigos empíricos. O contexto empírico da pesquisa é, portanto, uma informação importante para os pesquisadores e profissionais da indústria.
- A abordagem empírica utilizada no estudo define a terceira categoria da taxonomia. Segundo Smite et al. [SMI10], a pesquisa dentro de disciplinas imaturas tende a ser mais de natureza exploratória que a pesquisa em campos maduros que por sua vez, concentra-se mais no teste de hipóteses, métodos ou ferramentas. Os autores descrevem que a abordagem de um

estudo empírico pode ser classificada como uma avaliação empírica ou um estudo com base empírica. O estudo com base empírica sustenta suas conclusões em dados empíricos, mas não realiza nenhuma avaliação empírica. Por outro lado, se o estudo avalia uma prática, um método, um *framework* ou uma ferramenta, ele pode ser classificado como uma avaliação empírica.

- A descrição do método de pesquisa foi relatada pela maioria dos profissionais que participaram da *survey*, como sendo uma informação de contexto importante para compreensão dos resultados, definindo, portanto, a quarta categoria da taxonomia. Os principais elementos desta categoria foram extraídos do estudo conduzido por Easterbrook et al. [EAS08], que descreve cinco classes de métodos de pesquisa que são mais relevantes para pesquisa empírica em Engenharia de Software: experimento controlado, estudo de caso, *survey* e etnografia. Além destes métodos, incluiu-se também *grounded theory*. Glaser e Strauss [GLA09] descrevem em seu livro que é comum encontrar estudos empíricos, publicados em revistas, que apresentam resultados baseados em teoria fundamentada em dados (*grounded theory*).
- Os resultados da *survey* mostraram que, da mesma forma que o método de pesquisa é importante, a forma como os dados foram coletados também foi relacionado pelos respondentes como sendo uma informação necessária para interpretar os resultados e reproduzir o estudo em outro contexto. Portanto, a quinta categoria da taxonomia define a fonte de evidência empírica do estudo. Os elementos da categoria foram extraídos da revisão sistemática realizada por Smite et al. [SMI10], que mapeou os métodos mais usados por estudos empíricos em DDS, para coletar dados empíricos, tais como, observação, análise de documentos, entrevista, questionário e *post-mortem*.
- A sexta categoria compreende a avaliação dos resultados do artigo. A *survey* mostrou que essa informação é importante para os profissionais da indústria, uma vez que facilita o processo de busca por estudos empíricos em DDS. Os elementos da categoria foram extraídos da revisão sistemática realizada por Smite et al. [SMI10]. Os autores classificaram os resultados dos estudos em cinco categorias: história de sucesso, práticas bem

sucedidas, relatório de problemas, práticas fracassadas e história de falha. Incluiu-se também o elemento “resultados preliminares”, que caracteriza os estudos que apresentam resultados preliminares de uma pesquisa em andamento.

- A sétima categoria define o sujeito do estudo, que pode ser classificado como estudantes ou profissionais da indústria. Conforme evidenciado na *survey*, a identificação do sujeito da pesquisa ajuda na interpretação dos resultados.

O sujeito do estudo, definido na sétima categoria, pode ser complementado com informações adicionais de contexto, tais como, localização, distância geográfica e distância temporal dos profissionais da indústria. Desta forma, além de informar o sujeito do estudo, descreve-se também como os profissionais estão distribuídos neste cenário global de desenvolvimento de software. Este tipo de classificação é irrelevante para as pesquisas que utilizam estudantes como sujeitos do estudo. Portanto, para descrever a localização dos profissionais da indústria, foi acrescida à taxonomia proposta as categorias definidas por Smite et al. [SMI12]. A seguir, descrevemos resumidamente as quatro categorias finais da taxonomia e os seus elementos, que já foram apresentados no referencial teórico deste trabalho na seção 2.2.2:

- Localização: pode ser *onshore*, ou seja, dentro do mesmo país, ou *offshore*, quando é em outro país.
- Entidade Legal: pode ser *insourcing*, se o desenvolvimento ocorre em outro site da mesma empresa e *outsourcing*, se for com uma empresa diferente.
- Distância geográfica: quando for *onshore*, pode ser “próxima”, onde nenhum voo é necessário, ou “distante”, onde um voo é necessário. Para *offshore*, a classificação pode ser “perto”, onde o tempo de voo é menor que duas horas ou “longe”, no qual o tempo de voo é maior que duas horas.
- Diferença temporal: para situações *onshore*, pode ser “similar”, quando a diferença é igual ou menor que uma hora ou “diferente”, quando for maior que uma hora. Para o contexto *offshore*, podemos classificar como “pequeno”, quando a diferença de tempo for menor ou igual a quatro horas ou “grande”, quando a diferença é maior que quatro horas.

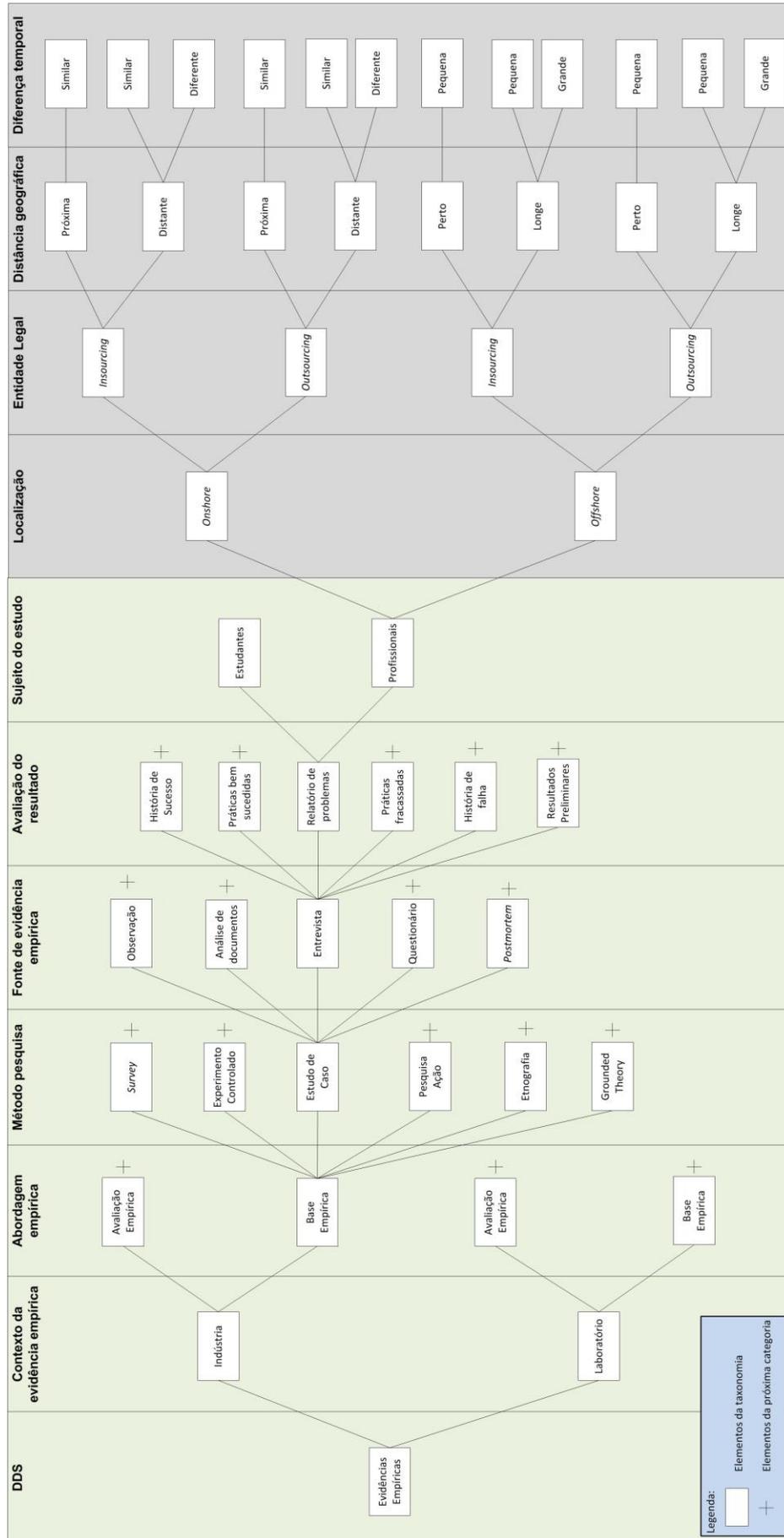


Figura 23 – Taxonomia proposta.

A taxonomia deve ser lida da esquerda para direita, começando pelo primeiro elemento que identifica os estudos que apresentam evidências empíricas, ilustrado na categoria “DDS”. Conforme indicado na legenda da Figura 23, os elementos de cada categoria são representados por um retângulo e possuem uma ligação direta com todos os elementos da próxima categoria. Os elementos que contêm um sinal de “+” podem se conectar com todos os elementos da próxima categoria e os elementos que não contêm esse sinal, não apresentam ligações válidas com a próxima categoria, como é o caso do elemento “Estudantes” da categoria “Sujeito do estudo”. As categorias que definem a localização da terceirização (*sourcing*) de um serviço ou produto não se aplicam quando o sujeito do estudo são estudantes.

Ao todo são 36 elementos distribuídos em 11 categorias que foram identificadas através da análise da base teórica, descrita no Capítulo 2, e dos resultados da *survey* apresentados no Capítulo 4. A Tabela 24 apresenta à síntese das categorias e os elementos de classificação propostos nesta pesquisa, assim como, a fonte onde foi identificado cada elemento.

Tabela 24- Categorias, elementos e fonte da taxonomia proposta.

<b>Categorias</b>	<b>Elementos</b>	<b>Fonte</b>
<b>1. DDS</b>	<ul style="list-style-type: none"> <li>▪ Evidência Empírica</li> </ul>	<i>Survey</i>
<b>2. Contexto da evidência empírica</b>	<ul style="list-style-type: none"> <li>▪ Laboratório</li> <li>▪ Indústria</li> </ul>	<i>Survey</i> e Base Teórica - [SMI10]
<b>3. Abordagem empírica</b>	<ul style="list-style-type: none"> <li>▪ Avaliação Empírica</li> <li>▪ Base Empírica</li> </ul>	Base Teórica - [SMI10]
<b>4. Método de pesquisa</b>	<ul style="list-style-type: none"> <li>▪ <i>Survey</i></li> <li>▪ Experimento Controlado</li> <li>▪ Estudo de Caso</li> <li>▪ Pesquisa Ação</li> <li>▪ Etnografia</li> <li>▪ <i>Grounded Theory</i></li> </ul>	<i>Survey</i> e Base Teórica - [EAS08], [GLA09]
<b>5. Fonte de evidência empírica</b>	<ul style="list-style-type: none"> <li>▪ Observação</li> <li>▪ Análise de documentos</li> <li>▪ Entrevista</li> <li>▪ Questionário</li> <li>▪ <i>Post-mortem</i></li> </ul>	<i>Survey</i> e Base Teórica - [SMI10]
<b>6. Avaliação do resultado</b>	<ul style="list-style-type: none"> <li>▪ História de sucesso</li> <li>▪ Práticas bem sucedidas</li> <li>▪ Relatório de problemas</li> <li>▪ Práticas fracassadas</li> </ul>	<i>Survey</i> e Base Teórica - [SMI10]

Categorias	Elementos	Fonte
	<ul style="list-style-type: none"> <li>▪ História de falha</li> <li>▪ Resultados preliminares</li> </ul>	
<b>7. Sujeito do estudo</b>	<ul style="list-style-type: none"> <li>▪ Estudantes</li> <li>▪ Profissionais</li> </ul>	Survey e Base Teórica - [SMI10]
<b>8. Localização</b>	<ul style="list-style-type: none"> <li>▪ <i>Onshore</i></li> <li>▪ <i>Offshore</i></li> </ul>	Base Teórica – [SMI12]
<b>9. Entidade legal</b>	<ul style="list-style-type: none"> <li>▪ <i>Insourcing</i></li> <li>▪ <i>Outsourcing</i></li> </ul>	Base Teórica – [SMI12]
<b>10. Distância geográfica</b>	<ul style="list-style-type: none"> <li>▪ Próxima (<i>Onshore</i>)</li> <li>▪ Distante (<i>Onshore</i>)</li> <li>▪ Perto (<i>Offshore</i>)</li> <li>▪ Longe (<i>Offshore</i>)</li> </ul>	Base Teórica – [SMI12]
<b>11. Distância temporal</b>	<ul style="list-style-type: none"> <li>▪ Similar (<i>Onshore</i>)</li> <li>▪ Diferente (<i>Onshore</i>)</li> <li>▪ Pequeno (<i>Offshore</i>)</li> <li>▪ Grande (<i>Offshore</i>)</li> </ul>	Base Teórica – [SMI12]

## 5.2 PROCESSO DE AVALIAÇÃO DA TAXONOMIA

A seguir serão apresentados os cenários de avaliação da taxonomia proposta, a seleção dos estudos, extração dos dados e os resultados desta avaliação.

### 5.2.1 Cenários

A taxonomia foi avaliada de acordo com quatro cenários, sendo três cenários de sucesso e um de falha. O primeiro cenário (C1) representa os artigos empíricos que foram classificados com sucesso na taxonomia, pois apresentaram todas as categorias de classificação do estudo. Os estudos empíricos com informações claras de contexto que não foram classificados na taxonomia (C2), caracterizaram falha no processo de avaliação, pois a taxonomia não apresentou os elementos necessários para classificar o estudo. Os estudos que não apresentaram informações claras de contexto (C3), não caracterizaram falha no processo de avaliação da taxonomia, uma vez que não foi possível determinar os elementos de classificação a partir da descrição de contexto do artigo. E por fim, o cenário no qual os artigos analisados não apresentaram evidências empíricas (C4), também caracterizaram sucesso, pois estudos não empíricos estão fora do escopo desta pesquisa. A Tabela 25 apresenta a descrição dos cenários, resultado esperado e o resultado da avaliação da taxonomia proposta.

Tabela 25 - Cenários de avaliação da taxonomia.

ID	Descrição do cenário	Resultado	Avaliação da taxonomia
C1	O artigo apresenta evidências empíricas e informações claras de contexto do estudo.	Artigo classificado na taxonomia.	Sucesso
C2	O artigo apresenta evidências empíricas e informações claras de contexto do estudo.	Artigo <b>não</b> classificado na taxonomia.	Falha
C3	O artigo apresenta evidências empíricas, mas não apresenta informações claras de contexto do estudo.	Artigo <b>não</b> classificado na taxonomia.	Sucesso
C4	O artigo não apresenta evidências empíricas.	Artigo <b>não</b> classificado na taxonomia.	Sucesso

### 5.2.2 Seleção dos estudos

A estratégia adotada para avaliação da taxonomia incluiu a seleção de artigos publicados na conferência ICGSE, uma vez que esta conferência é dedicada ao tema. Foram selecionados todos os artigos completos publicados na primeira e na oitava edição da conferência, que ocorreram em 2006 e 2013 respectivamente. Essas duas edições foram escolhidas por representarem períodos distintos de publicações nesta área, tornando o processo de avaliação da taxonomia mais completo. Os tutoriais, *short papers*, painéis e palestras foram excluídos da seleção. A Tabela 26 apresenta o número total de artigos selecionados em cada edição da ICGSE. A lista com todos os artigos analisados encontra-se no Apêndice D deste trabalho.

Tabela 26 – Conferência, ano e número de artigos.

Conferência	Ano	Número de artigos
1ª ICGSE	2006	21
8ª ICGSE	2013	25
	<b>Total</b>	<b>46</b>

### 5.2.3 Extração dos dados

Para a extração dos dados foi utilizado uma ficha de leitura desenvolvida no MS Excel. Os itens da ficha foram escolhidos conforme as categorias identificadas na taxonomia. Além das categorias, foi incluído um item para coletar o resultado do processo de avaliação da taxonomia. A seguir apresentamos os itens extraídos dos estudos:

- Identificação (ID)
- Referência bibliográfica
- Ano
- Evidências Empíricas
- Abordagem Empírica

- Contexto Empírico
- Método Pesquisa
- Fonte de evidência primária
- Outras fontes empíricas
- Resultado
- Sujeito do estudo
- Localização
- Entidade Legal
- Distância Geográfica
- Distância Temporal
- Resultado da avaliação (sucesso ou falha)

### 5.3 RESULTADOS DA AVALIAÇÃO

A seguir serão apresentados os resultados da avaliação da taxonomia por cenário e também será ilustrado um caso no qual um estudo pode ter múltiplas classificações.

#### 5.3.1 Resultados por cenário

Todos os artigos selecionados foram lidos integralmente e um conjunto de informações foi extraído de cada estudo, conforme descrito na seção 5.2.3. Dos 46 artigos selecionados, 26 apresentaram evidências empíricas, conforme ilustrado na Figura 24.

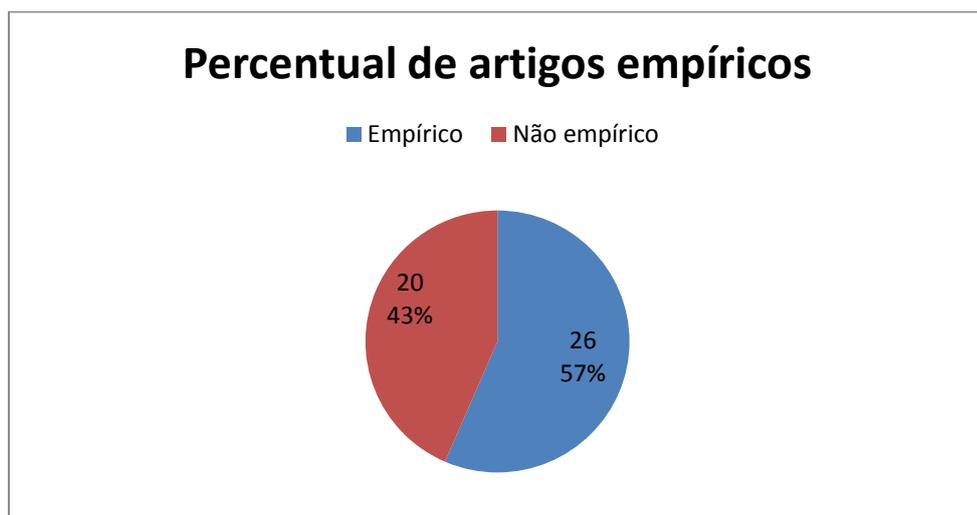


Figura 24 - Percentual de artigos empíricos.

A seguir apresentamos os resultados da avaliação da taxonomia proposta, conforme os cenários identificados.

### Cenário C1:

A Tabela 27 apresenta a lista de artigos classificados com sucesso na taxonomia proposta. De um total de 26 estudos empíricos, 17 artigos apresentaram informações claras de contexto. Os artigos A10, A11, A22 e A44 utilizaram estudantes como sujeitos do estudo, portanto, os elementos da taxonomia de Smite et al. [SMI12] não se aplicam neste contexto, conforme podemos ver nas células em cinza da Tabela 27. O mesmo é válido para o artigo A45, que utilizou profissionais como sujeitos do estudo, mas executou um experimento controlado com contexto empírico acadêmico.

Tabela 27 - Lista de artigos do cenário C1.

ID	Ano	Evidências Empíricas	Contexto Empírico	Abordagem Empírica	Método Pesquisa	Fonte evidência empírica	Resultado	Sujeito Estudo	Localização	Entidade Legal	Distância Geográfica	Distância Temporal
A1	2013	Sim	Indústria	Base Empírica	Estudo de Caso	Entrevistas	História de Sucesso	Profissionais	Offshore	Insourcing	Longe	Grande
A2	2006	Sim	Indústria	Avaliação Empírica	Estudo de Caso	Análise de docs	Práticas bem sucedidas	Profissionais	Offshore	Outsourcing	Longe	Grande
A3	2013	Sim	Indústria	Base Empírica	Estudo de Caso	Entrevistas	Práticas bem sucedidas	Profissionais	Offshore	Insourcing	Longe	Grande
A5	2006	Sim	Indústria	Base Empírica	Estudo de Caso	Observação	Relatório de Problemas	Profissionais	Offshore	Outsourcing	Longe	Grande
A6	2013	Sim	Indústria	Base Empírica	Grounded Theory	Entrevistas	História de Sucesso	Profissionais	Offshore	Outsourcing	Longe	Grande
A10	2006	Sim	Laboratório	Avaliação Empírica	Exp. Controlado	Questionário	História de Sucesso	Estudantes				
A11	2006	Sim	Laboratório	Avaliação Empírica	Exp. Controlado	Experimento	História de Sucesso	Estudantes				
A13	2006	Sim	Indústria	Base Empírica	Estudo de Caso	Observação	Relatório de Problemas	Profissionais	Offshore	Insourcing	Longe	Grande
A22	2013	Sim	Laboratório	Avaliação Empírica	Exp. Controlado	Experimento	Práticas bem sucedidas	Estudantes				
A24	2006	Sim	Indústria	Base Empírica	Estudo de Caso	Entrevistas	Relatório de Problemas	Profissionais	Offshore	Insourcing	Longe	Grande
A26	2013	Sim	Indústria	Base Empírica	Survey	Entrevistas	Relatório de Problemas	Profissionais	Offshore	Insourcing	Longe	Grande
A31	2013	Sim	Indústria	Base Empírica	Estudo de Caso	Entrevistas	Relatório de Problemas	Profissionais	Offshore	Insourcing	Longe	Grande
A36	2013	Sim	Indústria	Base Empírica	Estudo de Caso	Entrevistas	Relatório de Problemas	Profissionais	Offshore	Insourcing	Longe	Grande
A39	2013	Sim	Indústria	Base Empírica	Etnografia	Observação	Relatório de Problemas	Profissionais	Offshore	Outsourcing	Longe	Grande
A44	2013	Sim	Laboratório	Avaliação Empírica	Exp. Controlado	Questionário	História de sucesso	Estudantes				
A45	2013	Sim	Laboratório	Base Empírica	Exp. Controlado	Questionário	História de Sucesso	Profissionais				
A46	2013	Sim	Indústria	Base Empírica	Estudo de Caso	Entrevistas	História de Sucesso	Profissionais	Offshore	Outsourcing	Perto	Pequeno

### Cenário C2:

Todos os artigos que apresentaram evidências empíricas e informações claras de contexto foram classificados com sucesso na taxonomia proposta. Com relação ao conjunto de estudos selecionados, não encontramos nenhum artigo que caracterizasse falha da taxonomia. No entanto, entendemos que o resultado possa estar associado ao número de artigos analisados e também ao fato de termos usado apenas uma fonte para seleção dos artigos, a ICGSE.

### Cenário C3:

A Tabela 28 apresenta à lista dos nove artigos que não foram classificados na taxonomia devido à falta de clareza e, em alguns casos, a omissão de informações importantes sobre o contexto do estudo. Conforme descrito anteriormente, estes artigos não caracterizam falha no processo de validação da taxonomia, uma vez que não foi possível determinar os elementos de classificação a partir da descrição de contexto do

artigo. As células em vermelho da Tabela 28 representam as informações que não estão presentes no estudo ou não estão claras.

Tabela 28 - Lista de artigos do cenário C3.

ID	Ano	Evidências Empíricas	Contexto Empírico	Abordagem Empírica	Método Pesquisa	Fonte evidência empírica	Resultado	Sujeito Estudo	Localização	Entidade Legal	Distância Geográfica	Distância Temporal
A4	2006	Sim	Indústria	Avaliação Empírica	Estudo de Caso	Questionário	Resultados Preliminares	Profissionais				
A7	2013	Sim	Indústria	Base Empírica	Survey	Entrevistas	Relatório de Problemas	Profissionais				
A15	2013	Sim	Indústria	Base Empírica	Estudo de Caso	Questionário	Práticas bem sucedidas	Profissionais				
A16	2013	Sim	Indústria	Avaliação Empírica			Práticas bem sucedidas	Profissionais	Offshore	Outsourcing		
A17	2013	Sim	Indústria	Base Empírica	Grounded Theory	Entrevistas	Relatório de Problemas	Profissionais	Offshore			
A19	2013	Sim	Indústria	Base Empírica		Questionário	Resultados Preliminares	Profissionais				
A21	2006	Sim	Indústria	Base Empírica	Estudo de Caso	Questionário	Relatório de Problemas	Profissionais	Offshore			
A30	2013	Sim	Indústria	Avaliação Empírica		Entrevistas	Resultados preliminares	Profissionais				
A35	2006	Sim	Indústria	Base Empírica	Estudo de Caso	Entrevistas		Profissionais	Offshore	Insourcing	Longe	Grande

#### Cenário C4:

Por fim, a Tabela 29 lista os 20 artigos não empíricos que foram descartados, por não fazerem parte dos objetivos desta pesquisa. Assim como no cenário C3, esses artigos não caracterizam falha no processo de validação da taxonomia.

Tabela 29 - Lista de artigos do cenário C4.

ID	Ano	Evidências Empíricas
A8	2013	Não
A9	2013	Não
A12	2006	Não
A14	2013	Não
A18	2013	Não
A20	2006	Não
A23	2006	Não
A25	2006	Não
A27	2013	Não
A28	2006	Não
A29	2013	Não
A32	2006	Não
A33	2006	Não
A34	2006	Não
A37	2013	Não
A38	2006	Não
A40	2013	Não
A41	2006	Não
A42	2006	Não
A43	2006	Não

#### 5.3.2 Estudos com múltiplas classificações

A taxonomia também pode ser usada quando o estudo apresenta mais de uma classificação possível. Por exemplo, um estudo pode ter utilizado mais de um método de coleta de dados como fonte de evidência empírica ou os profissionais que participaram

como sujeitos do estudo podem estar distribuídos de diferentes formas (*onshore* ou *offshore*). A Tabela 30 apresenta um caso onde o artigo A3 apresenta três fontes de evidências empíricas. O estudo utilizou o método de entrevista como fonte primária e os métodos de observação e análise de documentos como fontes secundárias.

Tabela 30 - Artigo com múltiplas classificações (fonte de evidência empírica).

ID	Ano	Evidências Empíricas	Contexto Empírico	Abordagem Empírica	Método Pesquisa	Fonte evidência empírica	Resultado	Sujeito Estudo	Localização	Entidade Legal	Distância Geográfica	Distância Temporal
A3	2013	Sim	Indústria	Base Empírica	Estudo de Caso	Entrevistas	Práticas bem sucedidas	Profissionais	Offshore	Insourcing	Longe	Grande
A3	2013	Sim	Indústria	Base Empírica	Estudo de Caso	Observação	Práticas bem sucedidas	Profissionais	Offshore	Insourcing	Longe	Grande
A3	2013	Sim	Indústria	Base Empírica	Estudo de Caso	Análise de docs	Práticas bem sucedidas	Profissionais	Offshore	Insourcing	Longe	Grande

Da mesma forma, o artigo A6 pode ser classificado duas vezes na taxonomia de acordo com a localização dos profissionais terceirizados (*onshore* ou *offshore*), conforme ilustração da Tabela 31.

Tabela 31 - Artigo com múltiplas classificações (localização).

ID	Ano	Evidências Empíricas	Contexto Empírico	Abordagem Empírica	Método Pesquisa	Fonte evidência empírica	Resultado	Sujeito Estudo	Localização	Entidade Legal	Distância Geográfica	Distância Temporal
A6	2013	Sim	Indústria	Base Empírica	Grounded Theory	Entrevistas	História de Sucesso	Profissionais	Offshore	Outsourcing	Longe	Grande
A6	2013	Sim	Indústria	Base Empírica	Grounded Theory	Entrevistas	História de Sucesso	Profissionais	Onshore	Outsourcing	Distante	Similar

Portanto, ter uma classificação clara de cada estudo, simplifica o processo de busca por artigos descrevendo uma situação específica de contexto. A taxonomia é particularmente importante para os pesquisadores que estão à procura de trabalhos relacionados ou profissionais da indústria que buscam aprender com as pesquisas existentes relacionadas aos seus contextos ou áreas de interesse.

## 6 CONSIDERAÇÕES FINAIS

A partir dos avanços tecnológicos e da conversão de mercados locais em globais o DDS está cada vez mais presente no dia a dia das organizações de TI [PRI13], [BEE13]. Com o avanço da adoção do DDS nas organizações, pesquisadores da área de ES tem se defrontado com desafios e com a necessidade de propor novas soluções [GOM12]. Com a evolução recente da área, é preciso entender como os resultados empíricos gerados a partir de estudos científicos devem ser relatados para ser útil para profissionais na indústria e pesquisadores, sendo este o tema central de estudo desta pesquisa.

Os estudos realizados apresentam indícios de que a área de DDS necessita de mais pesquisas voltadas a esta mesma temática sob a perspectiva adotada nesta pesquisa: sistematização de evidências empíricas em DDS. Constatou-se que atualmente o processo de busca por estudos relacionados com DDS é custoso e muitas vezes não reflete o estado atual da área devido a grande diversidade de termos e informações de contexto envolvidas. Além disso, a terminologia aplicada ao DDS não é padronizada e muitas vezes o contexto de um estudo empírico não é descrito de forma clara, gerando, portanto, uma demanda para novos estudos de sistematização das evidências empíricas nesta área.

O objetivo desta dissertação, conforme apresentado na seção 1.1 deste volume foi alcançado com a realização de um estudo empírico, de natureza exploratória, contribuindo para a literatura e para a indústria ao fornecer uma taxonomia para compreender como as evidências empíricas em DDS estão sistematizadas e, portanto, tornar a busca por estudos empíricos um processo mais simples.

Este trabalho apresentou uma análise sobre as pesquisas científicas conduzidas na área de sistematização das evidências empíricas em DDS buscando oportunidades de pesquisa e um embasamento teórico para pesquisas futuras. Os trabalhos relacionados analisados forneceram uma visão de algumas das principais contribuições realizadas. Uma *survey* foi conduzida com especialistas que estão inseridos no ambiente de desenvolvimento distribuído e contribuíram com a identificação das informações de contexto mais relevantes em estudos com evidências empíricas em DDS.

A partir dos resultados obtidos nos estudos de base teórica, no mapeamento sistemático e na *survey* com especialistas, foi proposta uma taxonomia para a sistematização das evidências empíricas em DDS. A taxonomia visa auxiliar na compreensão de como as evidências empíricas geradas a partir de estudos científicos devem ser relatadas para ser útil para pesquisadores e profissionais na indústria. A

taxonomia proposta é uma tentativa de contribuir com o amadurecimento desta área de pesquisa, tornando mais fáceis o desenvolvimento de novos estudos empíricos.

## 6.1 CONTRIBUIÇÕES DA PESQUISA

A principal contribuição desta pesquisa é a proposta de uma taxonomia para sistematização das evidências empíricas em DDS (apresentado no Capítulo 5), que contempla elementos da literatura e a experiência de pesquisadores e profissionais da indústria que atuam com desenvolvimento distribuído de software. Para a academia o modelo proposto agrega conhecimentos, gerados a partir de estudos exploratórios, à Engenharia de Software ao apresentar uma proposta inicial de uma forma de sistematização das evidências empíricas ainda pouco explorada nesta área. A taxonomia apresentada serve também como um referencial teórico-prático para a indústria buscando melhorar a publicação de estudos empíricos e relatos de experiência, assim como, a busca por estudos empíricos em DDS.

Ao longo do processo de formulação da taxonomia proposta, descreveram-se as características da área de DDS, os desafios e estratégias de sistematização do conhecimento em ES (Capítulo 2), identificaram-se os tipos de informação de contexto que são mais relevantes para indústria e academia e as lições aprendidas que foram identificadas através da análise crítica dos resultados da *survey* (Capítulo 4). Adicionalmente, este estudo contribui com uma avaliação inicial da utilização da taxonomia proposta com estudos empíricos em DDS (Capítulo 5).

## 6.2 LIMITAÇÕES DA PESQUISA

Uma das principais limitações desta pesquisa está relacionada às restrições derivadas da metodologia de pesquisa adotada. Neste caso, o número de participantes entrevistados na parte empírica do estudo reduz a generalização dos resultados obtidos. Entretanto, os resultados da *survey* são sustentados na base teórica estudada, o que permite um bom grau de segurança nas conclusões obtidas. Isto também é típico do tipo de pesquisa desenvolvida, exploratória e de base qualitativa, permitindo o uso de inferências nas conclusões obtidas [PRI03].

Em relação à taxonomia proposta, esta não contempla todos os elementos de classificação dos estudos empíricos em DDS, entretanto retrata os principais atributos que foram verificados na análise dos resultados da *survey* e na base teórica desta pesquisa.

Além disso, esta taxonomia trata-se de uma proposta inicial para a sistematização das evidências empíricas no desenvolvimento distribuído de software. Por este motivo, é passível de modificações a partir de sua aplicação em novos contextos e melhor entendimento.

Por fim, a avaliação da taxonomia foi limitada aos estudos publicados na ICGSE e, portanto, é possível que outras bases de dados apresentem características de contexto não encontradas na amostra utilizada nesta pesquisa.

### **6.3 PESQUISAS FUTURAS**

Identifica-se que esta linha de pesquisa possui um grande potencial de crescimento, tanto a curto quanto em longo prazo. A parceria entre a academia e a indústria, permite criar condições para realização de *surveys* com especialistas, experimentos, validações de hipóteses e obtenção de conhecimentos, decorrentes de uma sinergia positiva entre os parceiros.

A taxonomia proposta, apresentada no Capítulo 5, deve ser testada e avaliada de forma mais profunda, tanto pela academia como pela indústria, em relação à classificação das evidências empíricas geradas a partir de estudos científicos em DDS. A partir do desenvolvimento de avaliações do modelo proposto, podem ser geradas novas atualizações da taxonomia, visando atender as melhorias identificadas. Indica-se também como pesquisa futura, a comparação com outras taxonomias na área de Engenharia de Software Experimental, uma vez que a taxonomia proposta contém também elementos mais abrangentes da área de ES.

Outra oportunidade de trabalho, apresentada no Capítulo 3 (fase 5 do desenho de pesquisa), seria a criação de uma base de evidências empíricas em DDS e uma ferramenta de apoio a esta base. A ferramenta de apoio deve ser desenvolvida para formulação de consultas a base e também servirá de apoio à criação da taxonomia. A base de evidências empíricas permitirá a criação de uma terminologia mais completa em DDS e principalmente, a categorização das evidências empíricas nesta área.

Além disso, entende-se que a taxonomia pode ser usada para ajudar a categorizar estudos empíricos obtidos por pesquisadores de outras áreas da Engenharia de Software, de modo que a aplicabilidade da evidência pode ser mais sistematicamente compreendida. A taxonomia deve servir como base para estudos futuros, auxiliando pesquisadores, que publicam ou sintetizam evidências empíricas, e profissionais que estão interessados em publicar casos empíricos.

## REFERÊNCIAS BIBLIOGRÁFICAS

- [ABR04] Abran, A.; Moore, J. W.; Bourque, P.; Dupuis, R. "Guide to the Software Engineering Body of Knowledge (SWEBOK)". IEEE Computer Society 2004 Guide, 2004, 346p.
- [ALI10] Ali, N. M.; Hosking, J.; Grundy, J. "A Taxonomy of Computer-Supported Critics". In: International Symposium on Information Technology, 2010, pp. 1152-1157.
- [AUD07] Audy, J.; Prikladnicki, R. "Desenvolvimento Distribuído de Software – Desenvolvimento de Software com Equipes Distribuídas". Campus, 2007, 232p.
- [BEE13] Beecham, S.; OLeary, P.; Richardson, I.; Baker, S.; Noll, J. "Who Are We Doing Global Software Engineering Research For?". In: 8th IEEE International Conference on Global Software Engineering, 2013, pp. 26-29.
- [CAR99] Carmel, E. "Global Software Teams - Collaborating Across Borders and Time Zones". EUA: Prentice Hall, 1999, 269p.
- [CHA99] Chandrasekaran, B.; Josephson, J. R.; Benjamins, V.R. "What Are Ontologies, and Why Do We Need Them?". IEEE Intelligent Systems & their applications, vol. 14, 1999, pp. 20-26.
- [CHA12] Chatzikonstantinou, G.; Kontogiannis, K. "Policy Modeling and Compliance Verification in Enterprise Software Systems: A Survey". In: 6th IEEE International Workshop on the Maintenance and Evolution of Service-Oriented and Cloud-Based Systems, 2012, pp. 27-36.
- [CHE12] Chen, H.; Ragsdell, G.; O'Brien, A.; Nunes, M. B. "A Proposed Model of Knowledge Management in the Software Industry Sector". In: 7th International Conference on Digital Information Management, 2012, pp. 291-296.
- [CRN02] Crnkovic, I.; Hnich, B.; Jonsson, T.; Kiziltan, Z. "Specification, Implementation, and Deployment of Components: Clarifying Common Terminology and Exploring Component-Based Relationships". Communications of the ACM, vol. 45, 2002, pp. 35-40.
- [DAV88] Davis, A. M. "A Taxonomy for the Early Stages of the Software Development Life Cycle". The Journal of Systems and Software, vol. 8, 1988, pp. 297-311.
- [DEL04] Delgado, N.; Gates, A. Q.; Roach, S. "A Taxonomy and Catalog of Runtime Software-Fault Monitoring Tools". IEEE Transactions on Software Engineering, vol. 30, 2004, pp. 859-872.
- [EAS08] Easterbrook, S. M.; Singer, J.; Storey, M.; Damian, D. "Selecting Empirical Methods for Software Engineering Research". In: Guide to Advanced Empirical Software Engineering, Springer, Heidelberg, 2008, pp. 285-311.
- [FIN03] Fink, A. "The Survey Handbook". Los Angeles: Sage Publications, 2003, 184p.

- [FOR08] Forward, A; Lethbridge, T. C. "A Taxonomy of Software Types to Facilitate Search and Evidence-Based Software Engineering". In: Conference of the Center for Advanced Studies, 2008.
- [FRE13] Frei, R.; McWilliam, R.; Derrick, B.; Purvis, A.; Tiwari, A.; Di Marzo, G. "Self-Healing and Self-Repairing Technologies." In: International Journal of Advanced Manufacturing Technology, 2013, pp. 1-29.
- [FRI05] Friedman T. L. "The world is flat: brief history of the 21st century". New York: Farrar, Straus and Giroux, 2005, 488p.
- [GFP12] Ghanam, Y.; Maurer, F.; Abrahamsson, P. "Making the Leap to a Software Platform Strategy: Issues and Challenges". Information and Software Technology, vol. 54, 2012, pp. 968-984.
- [GHA12] Ghazarian, A. "Characterization of Functional Software Requirements Space: The Law of Requirements Taxonomic Growth." In: 20th IEEE International Requirements Engineering Conference, 2012, pp. 241-250.
- [GIL10] Gil, A.C. "Como elaborar projetos de pesquisa". São Paulo: Atlas, 5ª Edição, 2010, 184p.
- [GLA09] Glaser, B. G; Strauss, A. L. "The Discovery of Grounded Theory: Strategies for Qualitative Research". Chicago: Aldine Transaction, 2009, 271p.
- [GOM12] Gomes, V.; Marczak, S. "Problems? We All Know We Have Them. Do We Have Solutions Too? A Literature Review on Problems and Their Solutions in Global Software Development". In: International Conference on Global Software Development, 2012, pp.154-158.
- [GRI06] Grimstad, S.; Jørgensen, M.; Moløkken-Østvold, K. "Software Effort Estimation Terminology: The Tower of Babel". Information and Software Technology, vol. 48, 2006, pp. 302-310.
- [GUM06] Gumm, D. C. "Distribution Dimensions in Software Development Projects: A Taxonomy." IEEE Software, vol. 23, 2006, pp. 45-51.
- [HER01] Herbsleb, J.; Moitra, D. "Global Software Development". IEEE Software, vol. 16, 2001, pp. 16-20.
- [HER07] Herbsleb, J. "Global Software Engineering: The Future of Socio-technical Coordination". In: 29th International Conference on Software Engineering, 2007, pp. 188-198.
- [HOF11] Hofer, C. N.; Karagiannis, G. "Cloud Computing Services: Taxonomy and Comparison". Journal of Internet Services and Applications, vol. 2, 2011, pp. 81-94.
- [HOL06] Holiday, O. J. "Para sistematizar experiências", Brasília: Ministério do Meio Ambiente, 2006, 128p.
- [HOS09] Hossain, E.; Babar M. A.; Pail, H. Y. "Using Scrum in Global Software Development: A Systematic Literature Review". In: 4th International Conference on Global Software Engineering, 2009, pp. 175–184.

- [JIM09] Jimenez, M.; Piattini, M.; Vizcaino, A. "Challenges and improvements in distributed software development: A systematic review". *Advances in Software Engineering*, 2009, pp. 1-16.
- [KAJ01] Kajko-Mattsson, M.; Westblom, U.; Forssander, S.; Andersson, G.; Medin, M.; Ebarasi, S.; Fahlgren, T.; Johansson, S.; Tornquist, S.; Holmgren, M. "Taxonomy of Problem Management Activities". In: *European Conference on Software Maintenance and Reengineering*, 2001, pp. 1-10.
- [KAR98] Karolak, D. W. "Global Software Development - Managing Virtual Teams and Environments". *IEEE Computer Society*, 1998, 159p.
- [KAS05] Kasunic, M. "Designing an effective survey". Technical report, handbook. *Software Engineering Institute, Carnegie Mellon University*, 2005.
- [KHW02] Khwaja, A.; Urban, J. E. "A Synthesis of Evaluation Criteria for Software Specifications and Specification Techniques." *International Journal of Software Engineering and Knowledge Engineering*, vol. 12, 2002, pp. 581-599.
- [KIT07] Kitchenham, B.; Charters, S. "Guidelines for performing Systematic Literature Reviews in Software Engineering", Technical Report, *Keele University and Durham University*, 2007, 65p.
- [KIT08] Kitchenham, B.; Pfleeger, S. "Personal Opinion Surveys". *Guide to Advanced Empirical Software Engineering*, Springer, London, 2008.
- [LAI00] Laitenberger, O.; Debaud, J. "Encompassing Life Cycle Centric Survey of Software Inspection". *Journal of Systems and Software*, vol. 50, 2000, pp. 5-31.
- [LAN08] Lane, M.; Agerfalk, P. "On the Suitability of Particular Software Development Roles to Global Software Development". In: *International Congress on Global Software Engineering*, 2008, pp. 3-12.
- [MAL11] Mala, M.; Çil, I. "A Taxonomy for Measuring Complexity in Agent-Based Systems". In: *2nd IEEE International Conference on Software Engineering and Service Science*, 2011, pp. 851-854.
- [MAN03] Mäntylä, M.; Vanhanen, J.; Lassenius, C. "A Taxonomy and an Initial Empirical Study of Bad Smells in Code". *Conference on Software Maintenance*, 2003, pp. 381-384.
- [MAR12] Marques, B.; Rodrigues, R.; Conte, T. "Systematic literature reviews in distributed software development: A tertiary study". In: *IEEE International Conference on Global Software Engineering*, 2012, pp. 134-143.
- [MEH00] Mehta, R.; Medvidovic, N.; Phadke, S. "Towards a Taxonomy of Software Connectors". In: *International Conference on Software Engineering*, 2000, pp. 178-187.
- [MEH13] Mehmood A.; Jawawi, A. "Aspect-Oriented Model-Driven Code Generation: A Systematic Mapping Study". *Information and Software Technology*, vol. 55, 2013, pp. 395-411.

- [MOE08] Moe, B.; Smite, D. "Understanding a lack of trust in global software teams: a multiple-case study". *Software Process Improvement and Practice*, vol. 13, 2008, pp. 217-231.
- [MON12] Monperrus, M.; Eichberg, M.; Tekes, E.; Mezini, M. "What should Developers be Aware of? an Empirical Study on the Directives of API Documentation". *Empirical Software Engineering*, vol. 17, 2012, pp. 703-737.
- [MOZ10] Mozzato, R.; Grzybovski, D. "Análise de Conteúdo como Técnica de Análise de Dados Qualitativos no Campo da Administração: Potencial e Desafios", *Revista de Administração Contemporânea*, vol. 15, 2010, pp. 731-747.
- [NOV10] Novak, J.; Krajnc, A.; Žontar, R. "Taxonomy of Static Code Analysis Tools". In: *33rd International Convention on Information and Communication Technology*, 2010, pp. 418-422.
- [NUR04] Nurmulliani, N.; Zowghi, D.; Fowell, S. "Analysis of Requirements Volatility during Software Development Life Cycle". In: *Australian Software Engineering Conference*, 2004, pp. 28-37.
- [OLU07] Oluyomi, A.; Karunasekera, S.; Sterling, L. "A Comprehensive View of Agent-Oriented Patterns." *Autonomous Agents and Multi-Agent Systems*, vol. 15, 2007, pp. 337-377.
- [PAA08] Paasivaara, M.; Durasiewicz, S.; Lassenius, C. "Using scrum in a globally distributed project: a case study". *Software Process: Improvement and Practice*, vol. 13, 2008, pp. 527-544.
- [PET08] Petersen, k.; Feldt, R.; Mujtaba, S.; Mattsson, M. "Systematic mapping studies in software engineering". In: *12th International Conference on Evaluation and Assessment in Software Engineering (EASE)*, Italy, 2008, pp. 68-77.
- [PET11] Petersen, K. "Measuring and Predicting Software Productivity: A Systematic Map and Review". *Information and Software Technology*, vol. 53, 2011, pp. 317-343.
- [PRE94] Preece, J.; Rombach, H. D. "A Taxonomy for Combining Software Engineering and Human-Computer Interaction Measurement Approaches: Towards a Common Framework." *International Journal of Human - Computer Studies*, vol. 41, 1994, pp. 553-583.
- [PRIE02] Prieto-Diaz, R. "A Faceted Approach to Building Ontologies". In: *21st International Conference on Conceptual Modeling-ER*, 2002, pp. 7-11.
- [PRI03] Prikladnicki, R. "MuNDDoS - Um modelo de referência para Desenvolvimento Distribuído de Software". *Dissertação de Mestrado, Programa de Pós-Graduação em Ciência da Computação, PUCRS*, 2003, 144p.
- [PRI10] Prikladnicki, R.; Audy, J. "Process Models in the Practice of distributed Software Development: A Systematic Review of the Literature". *Information and Software Technology*, vol. 52, 2010, pp. 779–791.

- [PRI13] Prikladnicki, R.; Boden, A.; Avram, G.; Souza, C.; Wulf, V. "Data collection in global software engineering research: learning from past experience". *Empirical Software Engineering*, vol. 19, 2013, pp. 822-856.
- [RAH12] Rahman, A.; Sahibuddin, S.; Ibrahim, S. "A Taxonomy Analysis for Multi-Model Process Improvement from the Context of Software Engineering Processes and Services." *International Journal of Digital Content Technology and its Applications*, vol. 6, 2012, pp. 56-65.
- [RAM12] Ramirez, A.; Jensen, A.; Cheng, B. "A Taxonomy of Uncertainty for Dynamically Adaptive Systems". In: *ICSE Workshop on Software Engineering for Adaptive and Self-Managing Systems*, 2012, pp. 99-108.
- [ROB12] Robillard, P.; Lavallée, M. "Software Team Processes: A Taxonomy". In: *International Conference on Software and System Process*, 2012, pp. 101-109.
- [ROM11] Romero, R.; Díez, D.; Montero, S.; Díaz, P.; Aedo, I. "Towards a Systematic Development of RIAs: Taxonomy for Classifying Rich-User-Interface Components". In: *Actas de las 16th Jornadas de Ingenieria del Software y Bases de Datos*, 2011, pp. 197-210.
- [SAK09] Sakarkar, G.; Shelke, N. "A New Classification Scheme for Autonomous Software Agent". In: *International Conference on Intelligent Agent and Multi-Agent Systems*, 2009, pp. 1-2.
- [SAV07] Savola, R. "Towards a Security Metrics Taxonomy for the Information and Communication Technology Industry". In: *2nd International Conference on Software Engineering Advances*, 2007.
- [SCO13] Scopus. "What does Scopus cover?" Capturado em: <http://www.info.sciverse.com/scopus/scopus-in-detail/facts>. Março, 2013.
- [SMI08] Smite, D.; Wohlin, C.; Feldt, R.; Gorschek, T. "Reporting empirical research in global software engineering: a Classification Scheme". In: *International Conference on Global Software Engineering*, 2008, pp. 173–181.
- [SMI10] Smite, D.; Wohlin C.; Feldt R.; Gorschek T. "Empirical Evidence in Global Software Engineering: A Systematic Review". *Empirical Software Engineering Journal*, vol. 15, 2010, pp. 91–118.
- [SMI12] Smite, D.; Wohlin, C.; Galviņa, Z.; Prikladnicki, R. "An Empirically Based Terminology and Taxonomy for Global Software Engineering". *Empirical Software Engineering*, vol. 19, 2012, pp. 105-153.
- [STR12] Strode, D.; Huff, S. "A Taxonomy of Dependencies in Agile Software Development". In: *23rd Australasian Conference on Information Systems*, 2012.
- [SUL04] Sulistio, A.; Yeo, C.; Buyya, R. "A Taxonomy of Computer-Based Simulations and its Mapping to Parallel and Distributed Systems Simulation Tools". *Software Practice and Experience*, vol. 34, 2004, pp. 653-673.
- [TAI11] Taivalsaari, A.; Mikkonen, T. "Objects in the Cloud may be Closer than they Appear: Towards a Taxonomy of Web-Based Software". In: *13th IEEE*

International Symposium on Web Systems Evolution, 2011, pp. 59-64.

- [TEC13] Techio, A.; Prikladnicki, R. "Desafios na Sistematização das Evidências Empíricas em Desenvolvimento Distribuído de Software". VII Workshop de Desenvolvimento Distribuído de Software, 2013, 2p.
- [UTT12] Utting, M.; Pretschner, A.; Legeard, B. "A Taxonomy of Model-Based Testing Approaches". Software Testing Verification and Reliability, vol. 22, 2012, pp. 297-312.
- [VAS11] Vasantrao, K. "Need to Understand Uncertainty in the System Development Modeling Process". In: International Conference on Communication Systems and Network Technologies, 2011, pp. 619-623.
- [YAN09] Yang, C.; Wang, H. "Towards a Classification of Aspects". In: 1st International Workshop on Education Technology and Computer Science, 2009, pp. 508-512.
- [YIN05] Yin, R. "Estudo de Caso Planejamento e Métodos". São Paulo: Bookman, 2005, 205p.
- [WAL13] Walia, G.; Carver, J. "Using Error Abstraction and Classification to Improve Requirement Quality: Conclusions from a Family of Four Empirical Studies". Empirical Software Engineering, vol. 18, 2013, pp. 625-658.
- [WIK14] Wikipedia. "Empirical evidence". Capturado em [http://en.wikipedia.org/wiki/Empirical\\_evidence](http://en.wikipedia.org/wiki/Empirical_evidence), Jan 2014.
- [WUY09] Wuyts, K.; Scandariato, R.; De Decker, B.; Joosen, W. "Linking Privacy Solutions to Developer Goals". In: International Conference on Availability, Reliability and Security, 2009, pp. 847-852.
- [ZAV97] Zave, P. "Classification of Research Efforts in Requirements Engineering". ACM Computing Surveys, vol. 29, 1997, pp. 315-321.

# APÊNDICE A

## Protocolo do Estudo de Mapeamento Sistemático

### 1. Formulação da Questão

#### 1.1. Questões de Pesquisa

A revisão tem a pretensão de responder as seguintes questões:

- QP1: Quais são as áreas da Engenharia de Software que apresentam estudos com o objetivo de sistematizar o conhecimento, através de estratégias como taxonomias, esquemas de classificação ou terminologias e quais são os principais estudos publicados na literatura?
- QP2: Quais são os desafios e os resultados ou contribuições relatadas nos estudos sobre a utilização de uma ou mais estratégias de sistematização tais como taxonomias, esquemas de classificação ou terminologias?

#### 1.2. Qualidade e Amplitude da Questão

##### 1.2.1. Palavras Chaves – Sinônimos

- *Software development*
- *Software engineering*
- *Taxonomy*
- *Classification scheme*
- *Terminology*

##### 1.2.2. População

A população alvo a ser pesquisada são estudos de sistematização em Engenharia de Software, portanto, os termos a serem listados é apresentado a seguir:

- *Software Development*
- *Software Engineering*

##### 1.2.3. Intervenção

Para garantir uma abrangência maior dentro da área de Engenharia de Software, não será utilizado nenhum termo de intervenção.

#### **1.2.4. Resultado**

Os resultados serão medidos por meio das estratégias de sistematização utilizadas nos estudos, estas são:

- *Taxonomy*
- *Classification scheme*
- *Terminology*

#### **1.2.5. Efeito**

Pretende-se identificar evidências empíricas de sistematização do conhecimento em ES e levantar novas iniciativas de pesquisas no assunto.

#### **1.2.6. Aplicação:**

Este mapeamento sistemático se aplica para pesquisadores da Academia que estejam desenvolvendo pesquisas científicas e necessitem de informações inerentes. Também pode vir a servir para profissionais de Mercado que possuem o interesse em informações importantes a respeito do assunto.

#### **1.2.7. Projetos Experimentais**

Não será utilizado nenhum método experimental.

#### **1.2.8. Artigos de Controle**

- *Grimstad, S., M. Jørgensen, and K. Moløkken-Østfold, 2006. "Software Effort Estimation Terminology: The Tower of Babel." Information and Software Technology, 48 (4): 302-310.*
- *Forward, A. and T. C. Lethbridge, 2008. "A Taxonomy of Software Types to Facilitate Search and Evidence-Based Software Engineering". Proceedings of the 2008 Conference of the Center for Advanced Studies, CASCON'08.*
- *Smite, D; Wohlin, C; Feldt, R, Gorschek, T. 2008. "Reporting empirical research in global software engineering: a Classification Scheme." In:*

*Proceedings of International Conference on Global Software Engineering ICG-SE 2008, 173–181.*

- *Smite, D; Wohlin, Claes; Galvina, Zanes; Prikladnicki, Rafael. 2012. “An Empirically Based Terminology and Taxonomy for Global Software Engineering”. Empirical Software Engineering: An International Journal. DOI: 10.1007/s10664-012-9217-9.*

## 2. Seleção de Fontes

### 2.1. Definição dos Critérios para Seleção de Fontes

- Artigos de conferência e periódicos;
- Base de dados atualizada;
- Estudos empíricos ou relatos de experiência contidos de forma clara.

### 2.2. Idioma dos Estudos

O idioma utilizado é o inglês pela quantidade de publicações em conferências e periódicos neste idioma.

### 2.3. Identificação das Fontes

#### 2.3.1. Método de Busca das Fontes

Pesquisas pelas máquinas de busca das bibliotecas que indexam pesquisa científica.

#### 2.3.2. *String* de Busca

População: sistematização do conhecimento em Engenharia de Software.

P:= (Software Development <or> Software Engineering)

Intervenção: Nenhuma.

Outcome: estratégia de sistematização utilizada.

O: = (taxonomy <or> classification scheme <or> terminology)

String de busca final: P <and> O

String de busca: (("software development" <or> "software engineering") <and> ("taxonomy" <or> "classification scheme" <or> "terminology"))

### 2.3.3. Lista de Fontes

A seguinte base de dados será utilizada:

- Scopus (<http://www.scopus.com/>)

## 3. Seleção de Estudos

### 3.1. Critérios de Inclusão

- Artigos que relatem ao menos uma das estratégias de sistematização do conhecimento no contexto de Engenharia de Software;
- Os artigos devem estar disponíveis na web;
- Os artigos devem apresentar textos completos dos estudos em formato eletrônico para leitura;
- Artigos publicados em conferências e periódicos;
- Estudos primários ou secundários;
- Artigos na área de Ciência da Computação.

### 3.2. Critérios de Exclusão

Os artigos serão excluídos da pesquisa, caso tiverem relação com estes critérios:

- Artigos que não sejam da área de Ciência da Computação;
- Artigos que não sejam em sua totalidade no idioma inglês;
- Artigos cujo foco central não é sistematização de uma determinada área, processo, ferramenta, sistema ou tecnologia;
- Estudos publicados pelo mesmo autor, sobre o mesmo tópico de pesquisa (somente estudo mais completo será selecionado e os demais serão excluídos);
- Estudos que expressam a opinião pessoal de alguém e que não dependem de trabalhos relacionados e metodologias de pesquisa;

- Estudos que apenas sugerem estratégias de sistematização, tais como, taxonomia, esquema de classificação e terminologia, porém não utilizam as mesmas no estudo em questão;
- Estudos que mencionam ou utilizam estratégias de sistematização existentes na literatura, porém não desenvolvem nenhum trabalho de melhoria ou adaptação das mesmas.
- Artigos curtos (*short papers*).

### **3.3. Procedimentos de Seleção de estudos**

A seleção dos trabalhos utilizados na revisão se deu por meio de três etapas:

- Identificação dos artigos obtidos nas *engines* de busca.
- Exclusão dos artigos baseada na leitura do título, abstract e palavras chave.
- Exclusão dos artigos baseada na leitura completa.

## **4. Referências Bibliográficas**

Kitchenham, B; Charters, S. "Guidelines for performing Systematic Literature Reviews in Software Engineering", Technical Report, Keele University and Durham University, 2007, 65p.

## APÊNDICE B

### Referências dos estudos selecionados no mapeamento sistemático

P01	Ali, N. M., J. Hosking, and J. Grundy. 2010. "A Taxonomy of Computer-Supported Critics". Proceedings 2010 International Symposium on Information Technology - System Development and Application and Knowledge Society, ITSIm'10, 1152-1157
P02	Chatzikonstantinou, G. and K. Kontogiannis. 2012. "Policy Modeling and Compliance Verification in Enterprise Software Systems: A Survey". IEEE 6th International Workshop on the Maintenance and Evolution of Service-Oriented and Cloud-Based Systems, MESOCA 2012, 27-36
P03	Chen, H., G. Ragsdell, A. O'Brien, and M. B. Nunes. 2012. "A Proposed Model of Knowledge Management in the Software Industry Sector" 7th International Conference on Digital Information Management, ICDIM 2012, 291-296
P04	Crnkovic, I., B. Hnich, T. Jonsson, and Z. Kiziltan. 2002. "Specification, Implementation, and Deployment of Components: Clarifying Common Terminology and Exploring Component-Based Relationships." Communications of the ACM 45 (10): 35-40
P05	Davis, A. M. 1988. "A Taxonomy for the Early Stages of the Software Development Life Cycle." The Journal of Systems and Software 8 (4): 297-311
P06	Delgado, N., A. Q. Gates, and S. Roach. 2004. "A Taxonomy and Catalog of Runtime Software-Fault Monitoring Tools." IEEE Transactions on Software Engineering 30 (12): 859-872.
P07	Forward, A. and T. C. Lethbridge. 2008. "A Taxonomy of Software Types to Facilitate Search and Evidence-Based Software Engineering". Proceedings of the 2008 Conference of the Center for Advanced Studies, CASCON'08
P08	Frei, R., R. McWilliam, B. Derrick, A. Purvis, A. Tiwari, and G. Di Marzo Serugendo. 2013. "Self-Healing and Self-Repairing Technologies." International Journal of Advanced Manufacturing Technology: 1-29
P09	Ghanam, Y., F. Maurer, and P. Abrahamsson. 2012. "Making the Leap to a Software Platform Strategy: Issues and Challenges." Information and Software Technology 54 (9): 968-984.
P10	Ghazarian, A. 2012. "Characterization of Functional Software Requirements Space: The Law of Requirements Taxonomic Growth." 20th IEEE International Requirements Engineering Conference, 241-250
P11	Grimstad, S., M. Jørgensen, and K. Moløkken-Østfold. 2006. "Software Effort Estimation Terminology: The Tower of Babel." Information and Software Technology 48 (4): 302-310
P12	Gumm, D. C. 2006. "Distribution Dimensions in Software Development Projects: A Taxonomy." IEEE Software 23 (5): 45-51
P13	Höfer, C. N. and G. Karagiannis. 2011. "Cloud Computing Services: Taxonomy and Comparison." Journal of Internet Services and Applications 2 (2): 81-94.
P14	Kajko-Mattsson, M., U. Westblom, S. Forssander, G. Andersson, M. Medin, S. Ebarasi, T. Fahlgren, S. -E Johansson, S. Törnquist, and M. Holmgren. 2001. "Taxonomy of Problem Management Activities". Proceedings of the European Conference on Software Maintenance and Reengineering, CSMR, 1-10
P15	Khwaja, A. and J. E. Urban. 2002. "A Synthesis of Evaluation Criteria for Software Specifications and Specification Techniques." International Journal of Software Engineering and Knowledge Engineering 12 (5): 581-599

P16	Laitenberger, O. and J. -M Debaud. 2000. "Encompassing Life Cycle Centric Survey of Software Inspection." <i>Journal of Systems and Software</i> 50 (1): 5-31
P17	Mala, M. and I. Çil. 2011. "A Taxonomy for Measuring Complexity in Agent-Based Systems". ICSESS 2011 - Proceedings: 2011 IEEE 2nd International Conference on Software Engineering and Service Science, 851-854
P18	Mäntylä, M., J. Vanhanen, and C. Lassenius. 2003. "A Taxonomy and an Initial Empirical Study of Bad Smells in Code". <i>Conference on Software Maintenance</i> , 381-384
P19	Mehmood A., Jawawi D.N.A., 2013. "Aspect-Oriented Model-Driven Code Generation: A Systematic Mapping Study." <i>Information and Software Technology</i> 55 (2): 395-411
P20	Mehta, Nikunj R., Nenad Medvidovic, and Sandeep Phadke. 2000. "Towards a Taxonomy of Software Connectors". <i>Proceedings - International Conference on Software Engineering</i> , 178-187
P21	Monperrus, M., M. Eichberg, E. Tekes, and M. Mezini. 2012. "What should Developers be Aware of? an Empirical Study on the Directives of API Documentation." <i>Empirical Software Engineering</i> 17 (6): 703-737
P22	Novak, J., A. Krajnc, and R. Žontar. 2010. "Taxonomy of Static Code Analysis Tools". <i>MIPRO 2010 - 33rd International Convention on Information and Communication Technology, Electronics and Microelectronics</i> , 418-422
P23	Nurmuliani, N., D. Zowghi, and S. Fowell. 2004. "Analysis of Requirements Volatility during Software Development Life Cycle". <i>Proceedings of the Australian Software Engineering Conference, ASWEC</i> , 28-37
P24	Oluyomi, A., S. Karunasekera, and L. Sterling. 2007. "A Comprehensive View of Agent-Oriented Patterns." <i>Autonomous Agents and Multi-Agent Systems</i> 15 (3): 337-377.
P25	Petersen, K. 2011. "Measuring and Predicting Software Productivity: A Systematic Map and Review." <i>Information and Software Technology</i> 53 (4): 317-343
P26	Preece, J. and H. D. Rombach. 1994. "A Taxonomy for Combining Software Engineering and Human-Computer Interaction Measurement Approaches: Towards a Common Framework." <i>International Journal of Human - Computer Studies</i> 41 (4): 553-583
P27	Rahman, A. A., S. Sahibuddin, and S. Ibrahim. 2012. "A Taxonomy Analysis for Multi-Model Process Improvement from the Context of Software Engineering Processes and Services." <i>International Journal of Digital Content Technology and its Applications</i> 6 (22): 56-65.
P28	Ramirez, A. J., A. C. Jensen, and B. H. C. Cheng. 2012. "A Taxonomy of Uncertainty for Dynamically Adaptive Systems". <i>ICSE Workshop on Software Engineering for Adaptive and Self-Managing Systems</i> , 99-108
P29	Robillard, P. N. and M. Lavallée. 2012. "Software Team Processes: A Taxonomy". <i>International Conference on Software and System Process, ICSSP 2012</i> - 101-109
P30	Romero, R., D. Díez, S. Montero, P. Díaz, and I. Aedo. 2011. "Towards a Systematic Development of RIAs: Taxonomy for Classifying Rich-User-Interface Components". <i>Actas de las 16th Jornadas de Ingenieria del Software y Bases de Datos, JISBD 2011</i> , 197-210
P31	Sakarkar, G. and N. M. Shelke. 2009. "A New Classification Scheme for Autonomous Software Agent". <i>International Conference on Intelligent Agent and Multi-Agent Systems, IAMA 2009</i> .

P32	Savola, R. 2007. "Towards a Security Metrics Taxonomy for the Information and Communication Technology Industry". 2nd International Conference on Software Engineering Advances - ICSEA 2007.
P33	Smite, D., Wohlin, C., R. Feldt, and Gorschek, T. 2008. "Reporting Empirical Research in Global Software Engineering: A Classification Scheme". Proceedings 3rd IEEE International Conference Global Software Engineering, ICGSE 2008, 173-181
P34	Smite, D., Wohlin, C., Galvina, Z., and Prikladnicki, R. 2012. "An Empirically Based Terminology and Taxonomy for Global Software Engineering." Empirical Software Engineering: 1-49.
P35	Strode, D. E. and S. L. Huff. , 2012. "A Taxonomy of Dependencies in Agile Software Development" Proceedings of the 23rd Australasian Conference on Information Systems
P36	Sulistio, A., C. S. Yeo, and R. Buyya. 2004. "A Taxonomy of Computer-Based Simulations and its Mapping to Parallel and Distributed Systems Simulation Tools." Software - Practice and Experience 34 (7): 653-673
P37	Taivalaari, A. and T. Mikkonen. 2011. "Objects in the Cloud may be Closer than they Appear: Towards a Taxonomy of Web-Based Software". Proceedings - 13th IEEE International Symposium on Web Systems Evolution, WSE 2011, 59-64
P38	Utting, M., A. Pretschner, and B. Legard. 2012. "A Taxonomy of Model-Based Testing Approaches." Software Testing Verification and Reliability 22 (5): 297-312.
P39	Vasantrao, K. V. 2011. "Need to Understand "Uncertainty" in the System Development Modeling Process". Proceedings - 2011 International Conference on Communication Systems and Network Technologies, CSNT 2011, 619-623
P40	Walia, G. S. and J. C. Carver. 2013. "Using Error Abstraction and Classification to Improve Requirement Quality: Conclusions from a Family of Four Empirical Studies." Empirical SE 18 (4): 625-658.
P41	Wuyts, K., R. Scandariato, B. De Decker, and W. Joosen. 2009. "Linking Privacy Solutions to Developer Goals". Proceedings - International Conference on Availability, Reliability and Security, ARES 2009, 847-852
P42	Yang, C. and H. Wang. 2009. "Towards a Classification of Aspects". Proceedings of the 1st International Workshop on Education Technology and Computer Science, ETCS 2009, 508-512
P43	Zave, Pamela. 1997. "Classification of Research Efforts in Requirements Engineering." ACM Computing Surveys 29 (4): 315-321

## APÊNDICE C

### PROTOCOLO DA SURVEY COM ESPECIALISTAS

#### Protocolo da Survey: Sistematização das Evidências Empíricas em Desenvolvimento Distribuído de Software

##### Objetivo

Identificar que tipo de informação de contexto é relevante, para profissionais da indústria e pesquisadores, em estudos com evidências empíricas em Desenvolvimento Distribuído de Software (DDS).

##### Característica-chave do método de pesquisa

Este é um roteiro para uma entrevista semiestruturada com questões abertas e fechadas.

##### Unidades de análise

Profissionais que trabalham e pesquisam na área de Desenvolvimento Distribuído de Software.

##### Organização desse Protocolo

O protocolo será organizado conforme segue:

#### 1. Procedimentos

<b>a. Levantamento das questões e estruturação do guia</b>	
Participantes	Antônio Techio
Data	Maio de 2013
Local	FACIN PUCRS – Faculdade de Informática da PUCRS
<b>b. Revisão do guia</b>	
Participantes	Prof. Dr. Rafael Prikkladnicki
Data	Junho de 2013
Local	PUCRS
<b>c. Validação de face e conteúdo</b>	
Participantes	Profa. Dr. Carolyn Seaman (University of Maryland Baltimore County)
Data	Julho de 2013
Local	PUCRS

<b>d. Pré-teste</b>	
Participantes	Bernardo Estácio - Doutorando
Data	Agosto de 2013
Local	PUCRS
<b>e. Aplicação das entrevistas - Academia</b>	
Participantes	Pesquisadores com experiência em DDS
Data	Setembro de 2013
Local	PUCRS (Porto Alegre) ou Skype (outras cidades ou países)
<b>f. Aplicação das entrevistas - Indústria</b>	
Participantes	Profissionais da indústria com experiência em DDS
Data	Outubro de 2013
Local	PUCRS (Porto Alegre) ou Skype (outras cidades ou países)

## 2. Escolha das pessoas entrevistadas

Profissionais da indústria e pesquisadores que possuem experiência na área de Desenvolvimento Distribuído de Software.

## 3. Outros recursos utilizados

- a. Recursos tecnológicos
  - i. Skype (para entrevistas remotas)
  - ii. Microsoft Excel (tabulação dos dados)
- b. Recursos materiais
  - i. Um gravador para gravar as entrevistas presenciais
  - ii. Papel e caneta

## 4. Modelo de estudo e dimensões da pesquisa

O esquema a seguir representa graficamente os principais aspectos focados no desenvolvimento deste trabalho. O esquema foi baseado do aprofundamento da base teórica e nos resultados de um mapeamento sistemático realizado pelos autores. Abaixo descrevemos cada etapa do modelo de estudo:

- **Aspectos de DDS:** indicam as diferentes áreas e características no contexto do DDS que influenciam na sistematização de estudos em DDS.
- **Desafios na sistematização do conhecimento:** conjunto de desafios descritos na literatura.
- **Soluções e estratégias para sistematização do conhecimento:** conjunto de soluções e estratégias utilizadas em diversos estudos para sistematização do conhecimento.
- **Sistematização das Evidências Empíricas em DDS:** tema central desta pesquisa.

- **Motivações:** explora o que motiva pesquisadores e profissionais da indústria a buscar estudos empíricos na literatura e o que eles esperam encontrar nesses estudos.
- **Expectativas:** explora o que os pesquisadores esperam encontrar em estudos com evidências empíricas em DDS.
- **Desafios:** desafios na busca por estudos com evidências empíricas em DDS.
- **Contexto:** busca identificar quais informações de contexto são relevantes para indústria e academia em estudos empíricos em DDS.

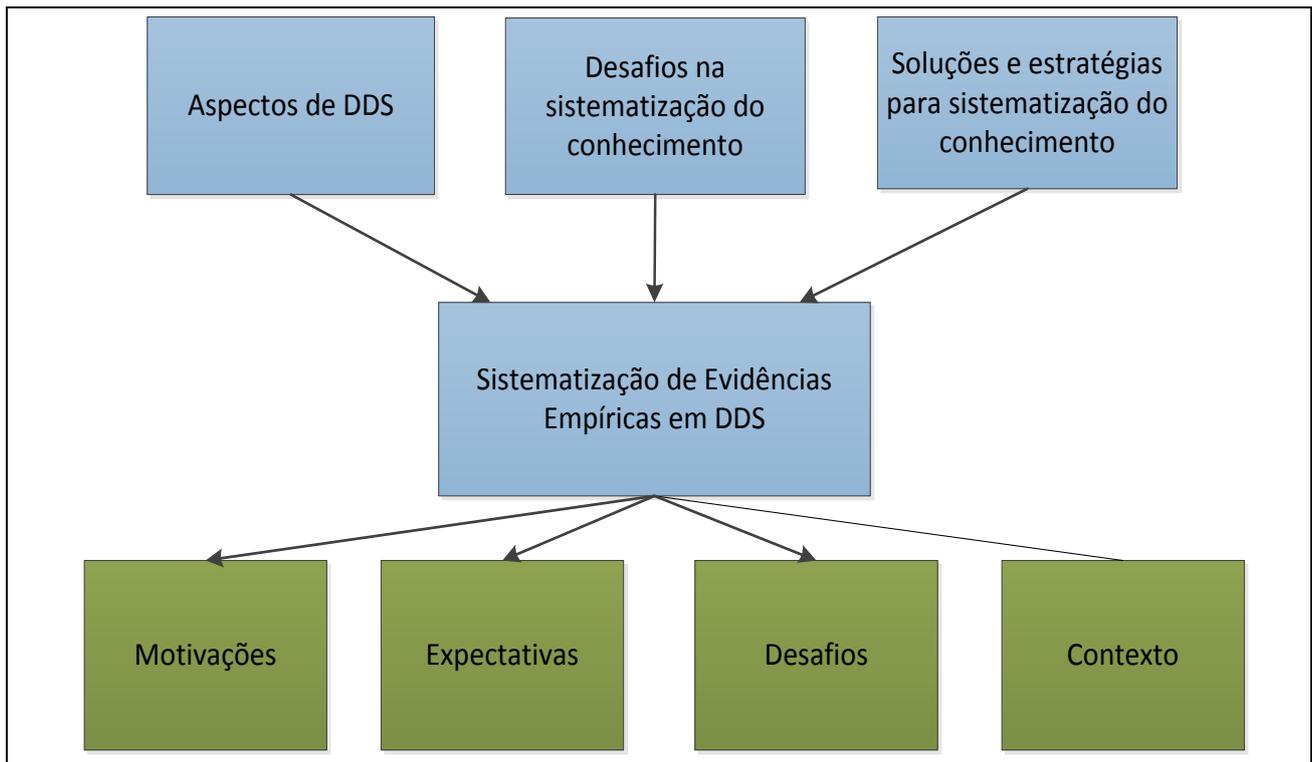


Figura 1 - Modelo do estudo e dimensões da pesquisa.

## 5. Coleta de dados

Entrevistas semiestruturadas com questões abertas e fechadas.

## 6. Análise de dados

Após a transcrição das entrevistas será realizada uma análise de conteúdo. Desta forma, todas as entrevistas serão gravadas, transcritas e analisadas. Para cada pergunta da entrevista serão identificadas categorias onde os dados serão codificados.

## 7. Seleção dos estudos empíricos do questionário

A seção do roteiro de entrevista que tem como objetivo investigar as informações de contexto dos estudos empíricos será ilustrado com dois estudos empíricos publicados na revista *Software Process Improvement and Practice*<sup>8</sup> em 2008. O primeiro estudo, [MOE08], trata sobre o tema confiança em DDS e o segundo artigo, [PAA08], trata sobre o tema Agile (Scrum) em DDS. As fontes dos artigos foram selecionadas do estudo [SMI12], no qual os autores executaram uma revisão sistemática e usaram os estudos para avaliação da taxonomia proposta. Os critérios utilizados para escolha dos artigos foram:

- Artigo de revista;
- Estudo empírico;
- Classificado na taxonomia proposta por [SMI12].

## 8. Questões do guia de entrevista – Versão em Português

### Dados demográficos dos respondentes:

- País onde você nasceu:
- País onde você trabalha:
- Ocupação principal:  Academia  Indústria
- Função:  Professor e/ou Pesquisador  Executivo  Gerente  Líder de projeto  Membro do time de projeto
- Experiência na indústria com projetos distribuídos (com equipes distribuídas que colaboram em um projeto de software comum, enquanto trabalham com fronteiras geográficas, temporais, culturais, políticas e organizacionais, para realizar uma tarefa interdependente):  
 Menos de 2 anos  De 2 a 4 anos  De 4 a 6 anos  De 6 a 10 anos  Mais de 10 anos  N/A
- Experiência como pesquisador(a) na área de Desenvolvimento Distribuído de Software:  
 Menos de 2 anos  De 2 a 4 anos  De 4 a 6 anos  De 6 a 10 anos  Mais de 10 anos  N/A
- Educação profissional (maior):  
 Graduação  MBA / Especialização  Mestrado  Doutorado  Pós-Doutorado

<sup>8</sup> [http://onlinelibrary.wiley.com/journal/10.1002/\(ISSN\)1099-1670](http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1099-1670)

### **Questões sobre evidências empíricas:**

O objetivo desta seção é identificar que tipo de evidência empírica em estudos de Desenvolvimento Distribuído de Software (DDS) é relevante para profissionais e pesquisadores nesta área.

#### **Motivações:**

1. Para qual finalidade você buscaria artigos empíricos<sup>9</sup> na literatura descrevendo estudos científicos ou relatos de experiência de outras empresas em DDS?

#### **Expectativas:**

2. O que você espera encontrar nesses estudos empíricos?

#### **Desafios:**

3. Imagine o seguinte cenário: você está procurando artigos para compreender melhor algum aspecto do DDS em sua própria empresa ou está procurando artigos relacionados com a sua própria pesquisa. Como você classificaria a complexidade<sup>10</sup> dessa tarefa?

Complexidade muito baixa  Complexidade baixa  Complexidade moderada  Complexidade alta  Complexidade muito alta

4. Por quê?

### **Questões sobre o contexto dos estudos:**

As questões desta seção buscam identificar que tipo de informação de contexto dos estudos em DDS é relevante para profissionais e pesquisadores desta área. Informações de contexto estão relacionadas com as características do estudo, tais como, método de pesquisa, coleta de dados, características da população estudada (número de participantes, perfil, etc.), características do objeto de estudo (indústria ou academia), nível de dispersão (mesma localização física, distância nacional, distância continental ou distância global), entre outros. Leia as sentenças

---

<sup>9</sup> Definição da palavra empírico do dicionário de Oxford: com base em, preocupado com, ou verificável por observação ou experiência ao invés de teoria ou lógica pura.

<sup>10</sup> Nesta pesquisa, o termo complexidade está relacionado com o processo de pesquisar e selecionar estudos com evidências empíricas na literatura e as dificuldades inerentes a esse processo, tais como, terminologia não padronizada, informações de contexto dos estudos incompletas ou volume de artigos encontrados que não fazem parte do objetivo da pesquisa.

abaixo, extraídas de dois estudos empíricos em DDS, e responda as perguntas. As sentenças foram extraídas de diferentes seções de cada estudo.

**5. Dadas as sentenças abaixo, selecione as informações de contexto que são mais importantes ou relevantes em um estudo empírico em DDS:**

ARTIGO A:

*[ ] The objective of this article is to understand the importance of trust in GSD by describing the key factors that lead to a lack of trust, and the effect that a lack of trust has on a GSD team;*

*[ ] The context for the research was the Latvian software development company LatSoftware, situated in Riga (Latvia in Europe);*

*[ ] We collected data from supply teams that participated in four globally distributed software projects run by LatSoftware;*

*[ ] We choose four global software development projects that all reported problems with trust.*

*[ ] We report on a multiple-case holistic study, in which we studied one phenomenon in several projects in one company;*

*[ ] We used multiple data sources: qualitative interviews with project participants, results from postmortem meetings, and project problem reports;*

*[ ] A postmortem meeting involves every team member. It focuses first on describing what went well and what did not work in the project, then on conducting a root-cause analysis of the main issues that are identified.*

*[ ] Using postmortem meetings, it was possible to identify the root causes of problems related to trust;*

*[ ] For the analysis, we relied mainly on qualitative interviews, because these provide a rich picture of the reasons for, and effects of, a lack of trust;*

*[ ] We describe four global projects run in the investigated software house, describing why trust was lacking, and the effects of this on each project.*

5.1. Porque as informações de contexto selecionadas são importantes para você?

5.2. Na sua opinião, existe alguma informação de contexto que está faltando ou não está claro neste estudo?

**6. Dadas as sentenças abaixo, selecione as informações de contexto que são mais importantes ou relevantes em um estudo empírico em DDS:**

ARTIGO B:

*[ ] This paper reports a case study on agile practices in a 40-person development organization distributed between Norway and Malaysia.*

*[ ] Based on seven interviews in the development organization, we describe how scrum practices were successfully applied.*

*[ ] We performed a systematic literature review on practices used in distributed agile software development projects.*

*[ ] The research presented in this paper is a single case study of a large distributed product program using an agile process.*

*[ ] The development work is distributed between two countries, Norway and Malaysia, where our case company has its own offices.*

*[ ] We collected data using semi-structured, open-ended interviews that were recorded.*

*[ ] The qualitative data analysis was done by one of the researchers. Data was analyzed by qualitative coding according to instructions by Miles and Huberman (1994).*

*[ ] The team size varies from two to nine persons. Some teams are distributed between Norway and Malaysia. Whether a team is distributed or not can vary between iterations.*

*[ ] The time difference between the two sites is seven hours during the winter and six hours during the summer.*

*[ ] The main contribution of this paper is a detailed description on how scrum practices were successfully applied to a mid-sized distributed software product development program.*

6.1. Porque as informações de contexto selecionadas são importantes para você?

6.2. Na sua opinião, existe alguma informação de contexto que está faltando ou não está claro neste estudo?

## 9. Questões do guia de entrevista – Versão em Inglês

### Demographic data

- Country where you were born:
- Country where you work:
- Principal occupation:  Academia  Industry
- Function:  Professor and/or Researcher  Executive  Manager  Project Leader  Member of a project team
- Years of experience in Industry with distributed projects (with distributed teams who collaborate on a common software project while working across geographic, temporal, cultural, political, and organizational boundaries to accomplish an interdependent task):  
 Less than 2 years  from 2 to 4 years  from 4 to 6 years  from 6 to 10 years  
 More than 10 years  N/A
- Years of experience as a researcher in DSD area:  
 Less than 2 years  from 2 to 4 years  from 4 to 6 years  from 6 to 10 years  
 More than 10 years  N/A
- Professional Education (highest):  
 Graduate  MBA / Specialization  Master  PhD  Post-PhD

### Empirical Evidence Questions

The purpose of this study is to identify what kind of empirical evidence in Distributed Software Development studies is relevant for practitioners and researchers in this area.

#### **Motivations:**

1. For what purposes might you want to look at empirical papers in the literature describing scientific studies or experience reports from other companies about DSD?

#### **Expectations:**

2. What do you expect to find in such empirical studies?

**Challenges:**

3. Given the following scenario: suppose you are searching for papers in order to better understand some aspect of DSD in your own company or searching for papers related to your own research. How would you classify the complexity of this task?

Very Low Complexity  Low Complexity  Moderate Complexity  High Complexity  Very High Complexity

4. Why?

**Study Context Questions**

The questions in this section aim to identify what type of context information of studies in DSD is relevant for practitioners and researchers in this area. Context information is related to the characteristics of the study such as research method, data collection, population (number of participants, profile, etc.), characteristics of the study object (industry or academia), level of dispersion (same physical location, national distance, continental distance or global distance), among many others. Read the context information, extracted from two empirical studies in DSD, and answer the questions. These sentences were extracted from different sessions of each paper.

5. **Given the sentences below, please select the context information that are more important or relevant in an empirical study in DSD:**

PAPER A:

The objective of this article is to understand the importance of trust in GSD by describing the key factors that lead to a lack of trust, and the effect that a lack of trust has on a GSD team;

The context for the research was the Latvian software development company LatSoftware, situated in Riga (Latvia in Europe);

We collected data from supply teams that participated in four globally distributed software projects run by LatSoftware;

We choose four global software development projects that all reported problems with trust.

*[ ] We report on a multiple-case holistic study, in which we studied one phenomenon in several projects in one company;*

*[ ] We used multiple data sources: qualitative interviews with project participants, results from postmortem meetings, and project problem reports;*

*[ ] A postmortem meeting involves every team member. It focuses first on describing what went well and what did not work in the project, then on conducting a root-cause analysis of the main issues that are identified;*

*[ ] Using postmortem meetings, it was possible to identify the root causes of problems related to trust;*

*[ ] For the analysis, we relied mainly on qualitative interviews, because these provide a rich picture of the reasons for, and effects of, a lack of trust;*

*[ ] We describe four global projects run in the investigated software house, describing why trust was lacking, and the effects of this on each project.*

5.1. *Why the selected study context information is important to you?*

5.2. *Is there any context information that is either missing or unclear to you?*

**6. Given the sentences below, please select the context information that are more important or relevant in an empirical study in DSD:**

*PAPER B:*

*[ ] This paper reports a case study on agile practices in a 40-person development organization distributed between Norway and Malaysia.*

*[ ] Based on seven interviews in the development organization, we describe how scrum practices were successfully applied.*

*[ ] We performed a systematic literature review on practices used in distributed agile software development projects.*

*[ ] The research presented in this paper is a single case study of a large distributed product program using an agile process.*

[ ] *The development work is distributed between two countries, Norway and Malaysia, where our case company has its own offices.*

[ ] *We collected data using semi-structured, open-ended interviews that were recorded.*

[ ] *The qualitative data analysis was done by one of the researchers. Data was analyzed by qualitative coding according to instructions by Miles and Huberman (1994).*

[ ] *The team size varies from two to nine persons. Some teams are distributed between Norway and Malaysia. Whether a team is distributed or not can vary between iterations.*

[ ] *The time difference between the two sites is seven hours during the winter and six hours during the summer.*

[ ] *The main contribution of this paper is a detailed description on how scrum practices were successfully applied to a mid-sized distributed software product development program.*

6.1. *Why the selected study context information is important to you?*

6.2. *Is there any context information that is either missing or unclear to you?*

## **10. Referências Bibliográficas**

- [MOE08] N. B. Moe and D. Smite, "Understanding a lack of trust in global software teams: a multiple-case study", *Software Process Improvement and Practice*, vol. 13, no. 3, p. 217, 2008.
- [PAA08] Paasivaara, Maria; Durasiewicz, Sandra; Lassenius, Casper. "Using scrum in a globally distributed project: a case study". *Software Process: Improvement and Practice* 13(6): 527-544, 2008.
- [SMI12] Smite, D., C. Wohlin, Z. Galviņa, and R. Prikladnicki. 2012. "An Empirically Based Terminology and Taxonomy for Global Software Engineering." *Empirical Software Engineering: An International Journal*, Published online 18 July 2012. DOI: 10.1007/s10664-012-9217-9.

## APÊNDICE D

### Referências dos estudos selecionados para avaliação da taxonomia

A1	Al-Ani, B.; Marczak, S.; Prikladnicki, R.; Redmiles, D., "Revisiting the Factors that Engender Trust of Global Systems Engineers," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.31,40, 26-29 Aug. 2013
A2	Andaloussi, B.S.; Braun, A., "A Test Specification Method for Software Interoperability Tests in Offshore Scenarios: A Case Study," Global Software Engineering, 2006. ICGSE '06. International Conference on , vol., no., pp.169,178, Oct. 2006
A3	Anh Nguyen-Duc; Cruzes, D.S., "Coordination of Software Development Teams across Organizational Boundary -- An Exploratory Study," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.216,225, 26-29 Aug. 2013
A4	Aranda, G.N.; Vizcaino, A.; Cechich, A.; Piattini, M., "Technology Selection to Improve Global Collaboration," Global Software Engineering, 2006. ICGSE '06. International Conference on , vol., no., pp.223,232, Oct. 2006
A5	Avritzer, A.; Ostrand, T.; Weyuker, E.J., "Experience Developing Software Using a Globally Distributed Workforce," Global Software Engineering, 2006. ICGSE '06. International Conference on , vol., no., pp.179,186, Oct. 2006
A6	Bass, J.M., "Agile Method Tailoring in Distributed Enterprises: Product Owner Teams," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.154,163, 26-29 Aug. 2013
A7	Beecham, S.; OLeary, P.; Richardson, I.; Baker, S.; Noll, J., "Who Are We Doing Global Software Engineering Research For?," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.41,50, 26-29 Aug. 2013
A8	Bhadauria, A.; Bhattacharjee, S.; Anandkumar, C.B.; Puthiyonnan, S., "Sustaining High Performance in an Offshore Team in Globally Distributed Development: A Success Story," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.120,123, 26-29 Aug. 2013
A9	Bhattacharjee, S.; Bhadauria, A.; Kumar, I.K.G., "Managing Product Delivery Quality in Distributed Software Development," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.164,167, 26-29 Aug. 2013
A10	Bruegge, B.; De Lucia, A.; Fasano, F.; Tortora, G., "Supporting Distributed Software Development with fine-grained Artefact Management," Global Software Engineering, 2006. ICGSE '06. International Conference on , vol., no., pp.213,222, Oct. 2006
A11	Bruegge, B.; Dutoit, Allen H.; Wolf, Timo, "Sysiphus: Enabling informal collaboration in global software development," Global Software Engineering, 2006. ICGSE '06. International Conference on , vol., no., pp.139,148, Oct. 2006
A12	Caprihan, G., "Managing Software Performance in the Globally Distributed Software Development Paradigm," Global Software Engineering, 2006. ICGSE '06. International Conference on , vol., no., pp.83,91, Oct. 2006
A13	Casey, V.; Richardson, I., "Project Management within Virtual Software Teams," Global Software Engineering, 2006. ICGSE '06. International Conference on , vol., no., pp.33,42, Oct. 2006

A14	Costa, C.; Murta, L., "Version Control in Distributed Software Development: A Systematic Mapping Study," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.90,99, 26-29 Aug. 2013
A15	Deshpande, S.; Beecham, S.; Richardson, I., "Using the PMBOK Guide to Frame GSD Coordination Strategies," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.188,196, 26-29 Aug. 2013
A16	Deuter, A., "Slicing the V-Model -- Reduced Effort, Higher Flexibility," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.1,10, 26-29 Aug. 2013
A17	Dorairaj, S.; Noble, J.; Allan, G., "Agile Software Development with Distributed Teams: Senior Management Support," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.197,205, 26-29 Aug. 2013
A18	Dubey, A.; Hudepohl, J., "Towards Global Deployment of Software Engineering Tools," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.129,133, 26-29 Aug. 2013
A19	Dullemond, K.; van Gameraen, B., "What Distributed Software Teams Need to Know and When: An Empirical Study," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.61,70, 26-29 Aug. 2013
A20	Egan, R.W.; Tremaine, M.; Fjermestad, J.; Milewski, A.; O'Sullivan, P., "Cultural Differences in Temporal Perceptions and its Application to Running Efficient Global Software Teams," Global Software Engineering, 2006. ICGSE '06. International Conference on , vol., no., pp.55,61, Oct. 2006
A21	Ehrlich, Kate; Chang, K., "Leveraging expertise in global software teams: Going outside boundaries," Global Software Engineering, 2006. ICGSE '06. International Conference on , vol., no., pp.149,158, Oct. 2006
A22	Estler, H.C.; Nordio, M.; Furia, C.A.; Meyer, B., "Collaborative Debugging," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.110,119, 26-29 Aug. 2013
A23	Heistracher, T.; Kurz, T.; Marcon, G.; Masuch, Claudius, "Collaborative Software Engineering with a Digital Ecosystem," Global Software Engineering, 2006. ICGSE '06. International Conference on , vol., no., pp.119,126, Oct. 2006
A24	Holmstrom, H.; Conchuir, E.O.; Agerfalk, P.J.; Fitzgerald, B., "Global Software Development Challenges: A Case Study on Temporal, Geographical and Socio-Cultural Distance," Global Software Engineering, 2006. ICGSE '06. International Conference on , vol., no., pp.3,11, Oct. 2006
A25	Hsieh, Y., "Culture and Shared Understanding in Distributed Requirements Engineering," Global Software Engineering, 2006. ICGSE '06. International Conference on , vol., no., pp.101,108, Oct. 2006
A26	Kiani, Z.U.R.; Mite, D.; Riaz, A., "Measuring Awareness in Cross-Team Collaborations -- Distance Matters," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.71,79, 26-29 Aug. 2013
A27	Kuhrmann, M.; Mendez Fernandez, D.; Grober, M., "Towards Artifact Models as Process Interfaces in Distributed Software Projects," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.11,20, 26-29 Aug. 2013
A28	Kuni, R.; Bhushan, N., "IT Application Assessment Model for Global Software Development," Global Software Engineering, 2006. ICGSE '06. International Conference on , vol., no., pp.92,100, Oct. 2006
A29	Manz, C.; Stupperich, M.; Reichert, M., "Towards Integrated Variant Management in Global Software Engineering: An Experience Report," Global Software

	Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.168,172, 26-29 Aug. 2013
A30	Marques, A.B.; Carvalho, J.R.; Rodrigues, R.; Conte, T.; Prikladnicki, R.; Marczak, S., "An Ontology for Task Allocation to Teams in Distributed Software Development," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.21,30, 26-29 Aug. 2013
A31	Modi, S.; Abbott, P.; Counsell, S., "Negotiating Common Ground in Distributed Agile Development: A Case Study Perspective," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.80,89, 26-29 Aug. 2013
A32	Morgan, R.; Maurer, F., "MasePlanner: A Card-Based Distributed Planning Tool for Agile Teams," Global Software Engineering, 2006. ICGSE '06. International Conference on , vol., no., pp.132,138, Oct. 2006
A33	N. Mullick; M. Bass; Z. Houda; Paulish Paulish; M. Cataldo, "Siemens Global Studio Project: Experiences Adopting an Integrated GSD Infrastructure," Global Software Engineering, 2006. ICGSE '06. International Conference on , vol., no., pp.203,212, Oct. 2006
A34	Narayanan, S.; Mazumder, S.; Raju R, "Success of Offshore Relationships: Engineering team structures," Global Software Engineering, 2006. ICGSE '06. International Conference on , vol., no., pp.73,82, Oct. 2006
A35	O'Conchuir, E.; Holmstrom, H.; Agerfalk, P.J.; Fitzgerald, B., "Exploring the Assumed Benefits of Global Software Development," Global Software Engineering, 2006. ICGSE '06. International Conference on , vol., no., pp.159,168, Oct. 2006
A36	Paasivaara, M.; Lassenius, C.; Heikkila, V.T.; Dikert, K.; Engblom, C., "Integrating Global Sites into the Lean and Agile Transformation at Ericsson," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.134,143, 26-29 Aug. 2013
A37	Pierce, K.; Ingram, C.; Bos, B.; Ribeiro, A., "Experience in Managing Requirements between Distributed Parties in a Research Project Context," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.124,128, 26-29 Aug. 2013
A38	Prikladnicki, R.; Audy, J.L.N.; Evaristo, R., "A Reference Model for Global Software Development: Findings from a Case Study," Global Software Engineering, 2006. ICGSE '06. International Conference on , vol., no., pp.18,28, Oct. 2006
A39	Shah, H.; Harrold, M.J., "Culture and Testing: What is the Relationship?," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.51,60, 26-29 Aug. 2013
A40	Singh, H.K.; Uppili, U.; Pavuluri, U., "Making Continual Code Quality Monitoring and Control Processes Work in a Global Delivery Organization: COSMOS," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.173,177, 26-29 Aug. 2013
A41	Spanjers, H.; ter Huurne, M.; Bendas, D.; Graaf, B.; Lormans, M.; Van Solingen, R., "Tool Support for Distributed Software Engineering," Global Software Engineering, 2006. ICGSE '06. International Conference on , vol., no., pp.187,198, Oct. 2006
A42	Stark, J.; Arlt, M.; Walker, D.H.T., "Outsourcing Decisions and Models - Some Practical Considerations for Large Organizations," Global Software Engineering, 2006. ICGSE '06. International Conference, vol., no., pp.12,17, Oct. 2006
A43	Suling Zhang; Tremaine, M.; Fjermestad, J.; Milewski, A.; O' Sullivan, P.,

	"Delegation in Virtual Team: the Moderating Effects of Team Maturity and Team Distance," Global Software Engineering, 2006. ICGSE '06. International Conference on , vol., no., pp.62,68, Oct. 2006
A44	Tell, P.; Babar, M.A.; Grundy, J., "A Preliminary User Evaluation of an Infrastructure to Support Activity-Based Computing in Global Software Development (ABC4GSD)," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.100,109, 26-29 Aug. 2013
A45	van Gameraen, B.; van Solingen, R.; Dullemond, K., "Auto-erecting Virtual Office Walls: A Controlled Experiment," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.206,215, 26-29 Aug. 2013
A46	Zieris, F.; Salinger, S., "Doing Scrum Rather Than Being Agile: A Case Study on Actual Nearshoring Practices," Global Software Engineering (ICGSE), 2013 IEEE 8th International Conference on , vol., no., pp.144,153, 26-29 Aug. 2013