

**UNIVERSIDADE FEDERAL DE SÃO CARLOS**

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**RS/AT: UMA ABORDAGEM PARA APLICAÇÃO  
DE ANÁLISE TEMÁTICA EM REVISÃO  
SISTEMÁTICA**

**RAFAEL ADABO GASTALDI**

**ORIENTADORA: PROFA. DRA. SANDRA CAMARGO PINTO FERRAZ  
FABBRI**

São Carlos – SP

Agosto/2016

**UNIVERSIDADE FEDERAL DE SÃO CARLOS**

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**RS/AT: UMA ABORDAGEM PARA APLICAÇÃO  
DE ANÁLISE TEMÁTICA EM REVISÃO  
SISTEMÁTICA**

**RAFAEL ADABO GASTALDI**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação, área de concentração: Engenharia de Software

Orientadora: Profa. Dra. Sandra Camargo Pinto Ferraz Fabbri

São Carlos – SP

Agosto/2016



UNIVERSIDADE FEDERAL DE SÃO CARLOS  
Centro de Ciências Exatas e de Tecnologia  
Programa de Pós-Graduação em Ciência da Computação

---

Folha de Aprovação

---

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a defesa de Dissertação de Mestrado do candidato Rafael Adabo Gastaldi, realizada em 04/08/2016.

*Sandra C.*

---

Profª Drª. Sandra Camargo Pinto Ferraz Fabbri  
(UFSCar)

*Fabiano*

---

Prof. Dr. Fabiano Cutigi Ferrari  
(UFSCar)

---

Profª. Drª. Tayana Uchôa Conte  
(UFAM)

Certifico que a sessão de defesa foi realizada com a participação à distância do membro Profª. Drª. Tayana Uchôa Conte. Depois das arguições e deliberações realizadas, o participante à distância está de acordo com o conteúdo do parecer da comissão examinadora redigido no relatório de defesa do aluno Rafael Adabo Gastaldi.

*Sandra C.*

---

Profª Drª. Sandra Camargo Pinto Ferraz Fabbri  
Coordenadora da Comissão Examinadora  
(UFSCar)

## AGRADECIMENTOS

A Profa. Sandra Camargo Pinto Ferraz Fabbri por ter me selecionado para ser seu aluno e me dar a honra de tê-la como orientadora. Por direcionar minhas descobertas, pela sua grande dedicação e incentivo para sempre buscar a qualidade em meu trabalho e por me auxiliar a crescer como pessoa e profissional ao longo desses anos de convívio.

Ao Professor André Di Thommazo, meu professor de graduação que se tornou companheiro de laboratório na pós-graduação. Sempre me auxiliou com seus conselhos e conhecimento para condução de minha pesquisa e construção de minha vida pessoal e profissional. Agradeço também a todos os meus outros professores do Instituto Federal de São Paulo - Campus São Carlos por contribuir em minha formação profissional.

A minha família, em especial ao meu pai James, minha mãe Cristina e minha avó Avelina. Os grandes motivos da pessoa que me tornei e a quem serei eternamente grato pelo amor que recebi e lições que aprendi.

A todos meus companheiros de LAPES. Em especial a Elis Hernandes, auxiliou-me muito no início da pesquisa sempre estando presente para auxiliar-me nas dúvidas; ao Ivan Ervolino, companheiro de todo dia no caminho da universidade; aos colegas de laboratório André Abade, Odair Souza, Bento Siqueira, Thiago Lacerda, Cleiton Silva e tantos outros pelo companheirismo e amizade nesses anos de convívio.

Aos professores e funcionários do Departamento de Computação da UFSCar.

A Deus por me dar calma e persistência quando necessário, dedicação para o trabalho e sabedoria para conduzi-lo.

Ao CNPq pelo, apoio financeiro.

## ABSTRACT

**Context:** Systematic reviews have been increasingly applied in computing area and are used to perform a broad literature search and generate the state of the art, or evidences on an issue. For this, primary studies are identified, selected and entirely read, leading to a hard work of reading and analysis. Thematic analysis is a way to analyze data qualitatively, through the identification of relevant information that are grouped into categories. **Aim:** Present the RS/AT approach that applies thematic analysis in the context of systematic reviews, with computational support, aiming to extract relevant data from the literature gathered in the review. **Method:** Once defined how to use thematic analysis in the systematic review process, the computational support was developed based on the StArt tool (which supports systematic reviews) and the Insight tool (which supports the qualitative analysis and, therefore, the thematic analysis). Experimental studies were conducted to evaluate the feasibility of the approach application and possible contributions to the data extraction of the systematic review process. **Results:** Results showed that (i) when using the RS/AT approach, less experienced researchers in the subject of the review managed to extract similar data to researchers with more experience; (ii) the amount of data extracted with the approach was higher than using the extraction form; and (iii) the researchers considered helpful the computational support developed. **Conclusions:** The results of the studies provide insights that the RS/AT approach is feasible in practice, it avoids rework during the data extraction and it leads to a greater amount of extracted data when compared with the use of data extraction form, which is the traditional approach to extract data in the literature review process.

**Keywords:** secondary study, systematic review, qualitative analysis, thematic analysis, software engineering, Start, Insight

## RESUMO

**Contexto:** Revisões sistemáticas têm sido aplicadas cada vez mais na área de computação e são utilizadas para efetuar uma ampla busca na literatura e gerar o estado da arte, ou as evidências sobre determinado assunto. Para isso, estudos primários são identificados, selecionados e lidos na íntegra, levando a um árduo trabalho de leitura e análise. Análise temática é uma forma de analisar dados qualitativamente, por meio da identificação de informações relevantes que são agrupadas em categorias. **Objetivo:** Apresentar a abordagem RS/AT, que aplica análise temática no contexto de revisões sistemáticas, com suporte computacional, para extrair os dados relevantes da literatura abordada na revisão. **Metodologia:** Uma vez definida a forma de utilização da análise temática no processo de revisão sistemática, o suporte computacional foi desenvolvido com base na ferramenta Start (que apoia revisões sistemáticas) e na ferramenta Insight (que apoia a análise qualitativa e portanto, a análise temática). Estudos experimentais foram realizados para avaliar a viabilidade de aplicação da abordagem e se havia contribuição para a extração dos dados no processo de revisão sistemática. **Resultados:** Os resultados obtidos mostraram que (i) ao usar a abordagem RS/AT, pesquisadores menos experientes no tema da revisão conseguiram extrair dados similares aos dos pesquisadores com mais experiência; (ii) a quantidade de dados extraídos com a abordagem foi maior do que com o formulário de extração; e (iii) os pesquisadores consideraram útil o suporte computacional. **Conclusões:** Com os resultados dos estudos tem-se indícios de que a abordagem RS/AT é viável de ser aplicada, que ela evita o retrabalho na extração de dados e leva à extração de uma quantidade maior de dados quando comparada com a abordagem tradicional de se extraírem dados da literatura por meio de formulários de extração.

**Palavras-chave:** estudo secundário, revisão sistemática, análise qualitativa de dados, análise temática, engenharia de software, Start, Insight

## LISTA DE FIGURAS

2.1	Desenvolvimento do protocolo da Revisão Sistemática . . . . .	21
2.2	Etapas da condução da Revisão Sistemática . . . . .	23
2.3	Etapas da Revisão Sistemática Iterativa - adaptado de Fabbri et al. (2013) . . . . .	24
2.4	Gerenciador de referências EndNote . . . . .	25
2.5	Gerenciador de referências Mendeley . . . . .	26
2.6	Ferramenta Eppi Reviewer . . . . .	26
2.7	Ferramenta Revis - retirado de Felizardo et al. (2011) . . . . .	27
2.8	Processo simplificado para utilização da StArt, adaptado de Zamboni et al. (2010) . . . . .	28
2.9	Organização da ferramenta StArt e definição do protocolo . . . . .	29
2.10	Adicionar nova seção de busca . . . . .	30
2.11	Etapa de seleção dos estudos . . . . .	30
2.12	Formulários de extração na StArt . . . . .	31
2.13	Visualização de dados para sumarização . . . . .	32
2.14	Ciclo da pesquisa qualitativa . . . . .	33
2.15	Análise temática - adaptada de Cruzes e Dyba (2011) . . . . .	40
2.16	Ferramenta Nvivo . . . . .	42
2.17	Ferramenta MaxQDA . . . . .	42
2.18	Ferramenta AtlasTi . . . . .	43
2.19	Organização da ferramenta Insight . . . . .	45
2.20	Visualização dos dados e codificação dos documentos . . . . .	46

2.21	Mineração de texto em conjunto com a visualização de dados . . . . .	46
2.22	Etapas de <i>Axial e Selective Coding</i> na ferramenta Insight . . . . .	47
2.23	Exportação dos relatórios da análise . . . . .	48
2.24	Áreas de estudo . . . . .	52
3.1	Fluxos das etapas de revisão sistemática e da análise temática . . . . .	62
3.2	Análise temática em revisão sistemática apoiada por suporte computacional . . . . .	64
3.3	Planejamento da revisão na StArt . . . . .	65
3.4	Nuvem de palavras . . . . .	66
3.5	Modelo de mapa mental . . . . .	67
3.6	Fluxo dos dados para exportação da StArt . . . . .	68
3.7	Fluxo dos dados para importação na Insight . . . . .	69
3.8	Funcionalidades adicionadas na ferramenta para todos os tipos de projeto . . . . .	70
3.9	Visualização do tipo WordCloud gerada pela Insight . . . . .	70
3.10	Estrutura de funcionamento da WordCloud . . . . .	71
3.11	Estrutura de funcionamento do botão Mine . . . . .	72
3.12	Funcionalidades exclusivas para projetos advindos da ferramenta StArt . . . . .	73
4.1	Respostas Q1 - Qual destes perfis se adequa ao seu? . . . . .	77
4.2	Respostas Q2 - Qual seu curso? . . . . .	78
4.3	Respostas Q3 - Conduz ou já conduziu estudos secundários? Se sim, quais tipos? . . . . .	78
4.4	Respostas Q4 - Qual seu grau de experiência em estudos secundários? . . . . .	79
4.5	Respostas Q7 - Conduz ou já conduziu análise qualitativa de dados? Se sim, utilizou alguma técnica? . . . . .	80
4.6	Respostas Q8 - Qual seu grau de experiência em análise qualitativa de dados? . . . . .	80
4.7	Respostas Q9 - Considera relevante a aplicação de análise qualitativa em estudos secundários? . . . . .	81
4.8	Extração dos dados conduzida por P3 na Insight . . . . .	87
4.9	Extração dos dados conduzida por P2 na StArt . . . . .	87

4.10	Modelo de mapa mental obtido da extração dos dados por P1 . . . . .	88
4.11	Modelo de mapa mental obtido da extração dos dados por P2 . . . . .	89
4.12	Modelo de mapa mental obtido da extração dos dados por P3 . . . . .	89
4.13	Dificuldade de utilização descrita por cada participante . . . . .	99
4.14	Dificuldade de utilização da ferramenta Insight . . . . .	101
4.15	Utilidade da ferramenta Insight para extração dos dados em revisão sistemática	102
4.16	Modelo de mapa mental obtido da extração dos dados por P1 . . . . .	102
4.17	Modelo de mapa mental obtido da extração dos dados por P2 . . . . .	103
4.18	Modelo de mapa mental obtido da extração dos dados por P3 . . . . .	103
4.19	Modelo de mapa mental obtido da extração dos dados por P4 . . . . .	103
4.20	Modelo de mapa mental obtido da extração dos dados por P5 . . . . .	103
4.21	Modelo de mapa mental obtido da extração dos dados por P6 . . . . .	104
4.22	Modelo de mapa mental obtido da extração dos dados por P7 . . . . .	104
4.23	Modelo de mapa mental obtido da extração dos dados por P8 . . . . .	104

## LISTA DE TABELAS

1.1	Características da pesquisa qualitativa e quantitativa (DIAS; FERNANDES, 2000)	14
2.1	Etapas da meta-etnografia . . . . .	39
2.2	Planejamento do mapeamento sistemático conduzido . . . . .	49
2.3	Países de publicação . . . . .	51
2.4	Ferramentas utilizadas . . . . .	53
4.1	Definição do <i>survey</i> para avaliação de potenciais usuários . . . . .	76
4.2	Definição do Estudo de Caso I . . . . .	84
4.3	Caracterização dos participante do estudo . . . . .	85
4.4	Dados de planejamento da revisão disponibilizados à P1 e P3 . . . . .	86
4.5	Definição do Experimento I . . . . .	93
4.6	Caracterização dos participantes do experimento . . . . .	96
4.7	Dados do planejamento da revisão disponibilizado aos participantes . . . . .	97
4.8	Atividades realizadas no experimento . . . . .	97
4.9	Padronização da nomenclatura dados extraídos pelos participantes . . . . .	105
4.10	<i>Recall</i> dos participantes . . . . .	106

# SUMÁRIO

<b>CAPÍTULO 1 – INTRODUÇÃO</b>	<b>12</b>
1.1 Contextualização . . . . .	12
1.2 Motivação e Definição do Problema . . . . .	15
1.3 Metodologia de Desenvolvimento do Trabalho . . . . .	16
1.4 Organização do Trabalho . . . . .	17
<b>CAPÍTULO 2 – FUNDAMENTAÇÃO</b>	<b>19</b>
2.1 Considerações iniciais . . . . .	19
2.2 Estudos secundários . . . . .	20
2.2.1 Revisão Sistemática . . . . .	20
2.2.2 Mapeamento sistemático . . . . .	23
2.2.3 Ferramentas para auxiliar a condução de Estudos secundários . . . . .	24
2.2.4 StArt . . . . .	28
2.3 Análise qualitativa . . . . .	31
2.3.1 Análise qualitativa em Engenharia de Software . . . . .	34
2.3.2 Técnicas para auxiliar a condução da análise qualitativa . . . . .	35
2.3.3 Ferramentas para auxiliar a análise qualitativa de dados . . . . .	41
2.3.4 Insight . . . . .	43
2.4 Mapeamento sistemático sobre utilização de análise qualitativa em revisões sistemáticas . . . . .	48
2.5 Considerações finais . . . . .	59

<b>CAPÍTULO 3 – ANÁLISE TEMÁTICA EM REVISÃO SISTEMÁTICA</b>	<b>60</b>
3.1 Considerações iniciais . . . . .	60
3.2 Caracterização e desenvolvimento da análise temática em conjunto com estudos secundários . . . . .	61
3.2.1 Caracterização da análise temática em conjunto com a revisão sistemática	61
3.2.2 Aplicando análise temática em revisão sistemática com a abordagem RS/AT . . . . .	63
3.3 Integração e melhorias das ferramentas . . . . .	67
3.4 Considerações finais . . . . .	73
<b>CAPÍTULO 4 – ESTUDOS CONDUZIDOS</b>	<b>74</b>
4.1 Considerações iniciais . . . . .	74
4.2 Survey para avaliação de perfil de usuários . . . . .	75
4.2.1 Definição . . . . .	75
4.2.2 Planejamento . . . . .	75
4.2.3 Operação . . . . .	76
4.2.4 Análise e interpretação . . . . .	81
4.2.5 Ameaças à validade . . . . .	82
4.3 Estudo Piloto . . . . .	83
4.3.1 Seleção do Contexto . . . . .	84
4.3.2 Seleção das Variáveis . . . . .	84
4.3.3 Projeto Experimental . . . . .	85
4.3.4 Instrumentação . . . . .	86
4.3.5 Preparação e Execução do Estudo de Caso . . . . .	86
4.3.6 Análise e Discussão dos Resultados . . . . .	88
4.3.7 Ameaças à validade . . . . .	92
4.4 Experimento . . . . .	93

4.4.1	Seleção do Contexto . . . . .	93
4.4.2	Seleção das Variáveis . . . . .	94
4.4.3	Projeto Experimental . . . . .	95
4.4.4	Instrumentação . . . . .	96
4.4.5	Preparação e Execução do Experimento . . . . .	96
4.4.6	Análise e Discussão dos Resultados . . . . .	98
4.4.7	Ameaças à validade . . . . .	106
4.5	Considerações finais . . . . .	107
<b>CAPÍTULO 5 – CONCLUSÃO</b>		<b>109</b>
5.1	Contribuições e limitações . . . . .	111
5.2	Trabalhos futuros . . . . .	112
<b>REFERÊNCIAS</b>		<b>114</b>
<b>CAPÍTULO A –QUESTIONÁRIO PARA AVALIAÇÃO DE PERFIL DE USUÁRIO EM ESTUDOS SECUNDÁRIOS E ANÁLISE QUALITATIVA DE DADOS</b>		<b>118</b>
<b>CAPÍTULO B –FORMULÁRIOS DE CARACTERIZAÇÃO PARA O EXPERIMENTO</b>		<b>121</b>
<b>CAPÍTULO C –FORMULÁRIOS DE AVALIAÇÃO DO EXPERIMENTO</b>		<b>123</b>
C.1	Formulário 1 - <i>Feedback</i> de realização do experimento . . . . .	123
C.2	Formulário 2 - Avaliação da ferramenta Insight . . . . .	124

# Capítulo 1

## INTRODUÇÃO

---

---

### 1.1 Contextualização

O termo "Engenharia de Software Baseada em Evidência"(ESBE) foi empregado por Kitchenham, Dyba e Jorgensen (2004) e caracteriza o pensamento de que as propostas realizadas nessa área, para serem adotadas pela comunidade, precisam gerar evidências de suas vantagens e desvantagens, para que as pessoas que as usem saibam avaliar se estão fazendo escolhas acertadas. Essa mesma concepção já tinha sido declarada por Basili, Selby e Hutchens (1986), quando os autores declararam que tudo o que é proposto na área, antes de ser utilizado, deve ser devidamente experimentado e caracterizado.

Em decorrência da construção de evidências, como sugerido por Basili, Selby e Hutchens (1986), a busca por evidências se torna necessária para que se consiga, de fato, fazer escolhas acertadas sobre as tecnologias propostas na literatura e para que seja possível, por exemplo, caracterizar lacunas que demandam mais investigações e pesquisas para suprir problemas ou deficiências daquilo que já está disponível.

Assim, tanto para construir evidências como também para coletar evidências sobre um determinado tópico de interesse, é empreendida a atividade de pesquisa. Cada etapa da atividade de pesquisa, desde a coleta do material de estudo até a análise e sumarização dos dados obtidos, deve ser realizada criteriosamente e de modo a transmitir as informações obtidas a outros pesquisadores.

Kitchenham e Charters (2007) categorizam os tipos de pesquisa em estudos primário, secundário e terciário, descritos a seguir:

- *Estudos primários*: Investigam uma questão de pesquisa específica. Exemplo: experi-

mentos, estudos de caso e pesquisas de opinião (*survey*).

- *Estudos secundários*: Revisam estudos primários relativos a uma (ou algumas) questão(ões) de pesquisa com o objetivo específico de integração/síntese de evidências relacionadas a essa(s) questão(ões) de pesquisa. Exemplo: Revisões e Mapeamentos Sistemáticos.
- *Estudos terciários*: São revisões de estudos secundários relacionados a uma mesma questão de pesquisa.

Estudos secundários - revisão sistemática e mapeamento sistemático - são importantes aliados para, respectivamente, caracterizar o estado da arte ou elaborar um mapeamento sobre o tópico de pesquisa a ser abordado por meio da análise de estudos primários. Os estudos secundários possuem como objetivo obter e sumarizar as evidências/dados a respeito de um tema de interesse e é parte necessária da pesquisa pois auxilia a obter evidências do campo em que o trabalho está inserido.

A revisão sistemática da literatura é um meio de identificar, avaliar e interpretar todas as pesquisas disponíveis relevantes para uma determinada questão de pesquisa, área temática ou fenômeno de interesse (KITCHENHAM, 2004). Através de passos estruturados, bem definidos e claros, a revisão sistemática é utilizada para efetuar uma pesquisa ampla na literatura para obtenção do estado da arte a respeito do tema proposto.

O mapeamento sistemático é projetado para prover uma visão mais ampla de um tópico de pesquisa, de modo a estabelecer se há evidência de pesquisa nesse tópico e prover uma indicação da quantidade de evidência. O processo do mapeamento sistemático é similar ao da revisão sistemática. Kitchenham e Charters (2007) ressaltam que, se durante o exame inicial de um tópico (antes de fazer uma revisão sistemática), for constatado que o tópico é muito abrangente, um mapeamento sistemático pode ser mais indicado do que uma revisão sistemática.

Conduzir estudos secundários é uma prática quase que imprescindível para realização de qualquer pesquisa, pois eles auxiliam na obtenção do estado da arte do tema, isto é, da evidências sobre o tema a ser tratado. Para chegar ao estado da arte do tema tratado, a etapa de análise e obtenção dos dados deve ser realizada de forma cuidadosa e criteriosa.

Independentemente da categoria do estudo (primário, secundário ou terciário) a ser conduzido, a pesquisa irá lidar com dados e abordagens qualitativas e quantitativas, existindo claras diferenças entre elas (Tabela 1.1), e que deverão ser analisados de maneira adequada. A utilização da abordagem qualitativa ou quantitativa depende do tipo da pesquisa a ser conduzida, questões

que se pretende responder e objetivos a serem alcançados. Coleman e O'Connor (2007) ressaltam que enquanto estudos quantitativos possuem foco em questões como "Quanto?" e "Com que frequência?", estudos qualitativos focam em questões relativas a "Por quê?", "Como?" e "De que modo?".

**Tabela 1.1: Características da pesquisa qualitativa e quantitativa (DIAS; FERNANDES, 2000)**

	<b>QUANTITATIVA</b>	<b>QUALITATIVA</b>
<b>Paradigma</b>	Hipotético - dedutivo	Holístico - interpretativo
<b>Dados</b>	Representados numericamente Quantitativos Estruturados e não valorativos	Representados verbalmente Qualitativos Com maior riqueza de detalhes
<b>Papel do Pesquisador</b>	Observador Distância objetiva	Interpretador da realidade Imerso no contexto
<b>Abordagem</b>	Positivista Experimental Estudos confirmatórios	Interpretativa Não experimental Estudos exploratórios
<b>Análise</b>	Estatística Inferências a partir de amostras Teste de hipótese e teorias	Conteúdo ou caso Padrões a partir dos dados Hermenêutica e fenomenologia

Apesar das claras diferenças entre as abordagens, os métodos quantitativos e qualitativos não se excluem. Pope e Mays (1995) destacam que embora difiram quanto à forma e à ênfase, os métodos qualitativos trazem como contribuição ao trabalho de pesquisa uma mistura de procedimentos de cunho racional e intuitivo capazes de contribuir para a melhor compreensão dos fenômenos. Pode-se distinguir o enfoque qualitativo do quantitativo, mas não seria correto afirmar que guardam relação de oposição .

Assim como a revisão sistemática e o mapeamento sistemático, a análise qualitativa de dados também vem sendo mais utilizada pelas ciências humanas e biológicas, no entanto, principalmente na área de Engenharia de Software têm sido utilizadas técnicas de análise qualitativa para análise de dados provenientes de estudos de caso, como podem ser encontrados nos estudos de Travassos (2002) e Stol, Ralph e Fitzgerald (2016) no qual avaliam técnicas de análise qualitativa no contexto da área de Engenharia de Software.

Considerando o panorama apresentado, o desenvolvimento de métodos, abordagens e ferramentas para auxiliar a análise dos dados - principalmente análise qualitativa de dados - em revisões sistemáticas, que têm por objetivos construir evidências e, portanto, contribuem para a área de ESBE, torna-se relevante e necessário. A análise de dados qualitativos pode resultar em evidências com perspectivas diferentes em relação às análises puramente quantitativas. Nesse contexto, este trabalho apresenta a abordagem RS/AT que utiliza análise temática (uma técnica de análise qualitativa) em revisões sistemáticas, para auxiliar a extração de dados e construírem-

se evidências sobre o tópico investigado.

## 1.2 Motivação e Definição do Problema

Conforme visto na seção anterior, construir e coletar evidências é uma atividade relevante para a área de engenharia de software. Em particular, para a coleta de evidências, estudos secundários têm sido empregados e análise qualitativa têm sido utilizada com mais frequência em decorrência dos estudos experimentais realizados para se construírem as evidências.

Em particular, a revisão sistemática demanda tempo, dedicação e requer habilidade para interpretar os resultados coletados. Quando os resultados são dados qualitativos, em geral, esses dados trazem informações mais ricas, mas devem ser interpretados e sumarizados por métodos qualitativos. Seaman (2008) ressalta que a principal vantagem da utilização de métodos qualitativos é que forçam o pesquisador a se aprofundar na complexidade do problema, permitindo que a riqueza trazida pelos dados seja alcançada. Em contrapartida, essa autora também coloca que a análise qualitativa dos dados não é uma tarefa simples de ser conduzida, pode ser algumas vezes chata, muitas vezes tediosa e consumir mais tempo do que o esperado (SEAMAN, 2008).

Apesar das afirmações a respeito da complexidade e esforço necessário para realizar revisões sistemáticas e análise qualitativa de dados, vale ressaltar que ambos são importantes instrumentos no contexto de ESBE. Dixon-Woods et al. (2005) ressaltam que questões complexas exigem formas complexas de evidenciar. Os métodos atuais para a síntese de evidências têm, no entanto, tendido a favorecer formas quantitativas de análise e, muitas vezes, omitem as evidências qualitativas, podendo provocar perda de dados relevantes para a pesquisa.

Também no contexto da pesquisa qualitativa, Teixeira (2003) ressalta que pesquisadores que adotam as abordagens qualitativas têm mostrado preocupação com o rigor com que pretendem conduzir sua investigação. Para tanto buscam definir procedimentos que maximizam a confiabilidade, assim, os computadores se tornaram auxiliares importantes no manuseio e na análise de dados qualitativos. Apesar disso, nenhum sistema disponível pode substituir as qualidades interpretativas do pesquisador.

As afirmações apresentadas são reforçadas pelos comentários de Cruzes e Dyba (2011), que citam que o desenvolvimento de métodos de síntese para revisões sistemáticas trazem vantagens para elaboração de teorias relevantes para a prática e pesquisa na área de Engenharia de Software.

A análise de dados qualitativos pode resultar em evidências com perspectivas diferentes em

relação às análises puramente quantitativas. Nesse contexto, observa-se que o desenvolvimento de métodos, ferramentas e abordagens para guiar a utilização das técnicas de análise temática e revisão sistemática em conjunto são importantes.

Assim, o objetivo deste trabalho foi definir a abordagem RS/AT que utiliza a técnica de análise temática (uma técnica de análise qualitativa) em conjunto com revisão sistemática, apoiada com recursos computacionais. Para que esse objetivo fosse atingido, várias atividades foram realizadas, as quais compuseram a metodologia de desenvolvimento deste trabalho, como explicado na próxima seção.

### 1.3 Metodologia de Desenvolvimento do Trabalho

Com base no objetivo determinado anteriormente, estabeleceram-se algumas questões a serem investigadas no decorrer do trabalho, as quais consistiram em atividades que foram realizadas durante seu desenvolvimento. São elas:

1. *Quais etapas da revisão sistemática a análise qualitativa pode auxiliar?*

A revisão sistemática é um processo estruturado e com passos bem definidos. Já a análise qualitativa possui caráter exploratório e investigativo, sem necessariamente a definição de passos detalhados para sua condução. Assim, para abordar o ponto levantado na questão, investigou-se por meio de um mapeamento sistemático (descrito no Capítulo 2), qual técnica de análise qualitativa tem sido mais usada na condução de revisões sistemáticas e como ela tem sido usada nesse contexto.

2. *Quando utilizadas as técnicas de análise qualitativa e revisão sistemática em conjunto ainda se faz necessário utilizar o formulário de extração dos dados?*

Os formulários de extração são elaborados na fase de planejamento da revisão sistemática para guiar a busca por informações nos estudos primários e auxiliar a recuperação dos dados para a etapa de sumarização da revisão sistemática. Uma premissa deste trabalho foi que, utilizando-se análise temática para extrair as informações relevantes dos estudos primários, o formulário de extração de dados não precisaria ser elaborado durante o planejamento da revisão sistemática, pois as informações relevantes seriam obtidas conforme o processo de análise temática. Nesse sentido, foram conduzidos estudos para verificar se essa premissa era verdadeira.

3. *As ferramentas StArt e Insight, desenvolvidas por outros trabalhos deste grupo de pesquisa são passíveis de uma utilização conjunta para dar suporte à utilização da análise*

*temática na condução de revisão sistemática?*

Considerando que já havia no grupo de pesquisa uma ferramenta que dá suporte à condução de revisão sistemática - a ferramenta StArt (ZAMBONI et al., 2010) - e uma ferramenta que dá suporte à condução de análise temática - a ferramenta Insight (HERNANDES, 2014). Uma das atividades realizadas durante este trabalho foi fazer com que essas duas ferramentas pudessem atuar em conjunto para dar suporte à abordagem RS/AT. Para conduzir essa atividade com mais conhecimento, foram investigadas outras ferramentas de análise qualitativa e de revisão sistemática, de forma a extrair ideias que pudessem ajudar na integração da StArt com a Insight. Essa investigação está comentada no Capítulo 2.

4. *Quais vantagens e desvantagens podem ser observadas quando do uso da abordagem RS/AT em comparação com a condução da revisão sistemática de forma tradicional, isto é, com a extração de dados baseada no formulário de extração?*

Seaman (2008) cita que a principal vantagem de se utilizar métodos qualitativos é que forçam o pesquisador a explorar a complexidade do problema tratado e, portanto, os resultados são mais ricos e informativos em relação a dados puramente quantitativos. Porém, a mesma autora ressalta que a utilização de métodos qualitativos consomem mais tempo e são, muitas vezes, tediosos para serem aplicados. Nesse sentido, com base nos estudos conduzidos, analisaram-se os resultados com essa perspectiva de estabelecerem-se vantagens e desvantagem caso houvesse. Essas análises estão descritas no Capítulo 4, que relata os estudos realizados.

## 1.4 Organização do Trabalho

Esta trabalho está organizado em 5 capítulos. Este capítulo apresentou o contexto no qual esta dissertação está inserida, motivação e objetivos do trabalho bem como a metodologia de desenvolvimento do mesmo.

No capítulo 2 é apresentada a fundamentação teórica do trabalho baseada na revisão bibliográfica realizada. São abordados os conceitos relativos a estudos secundários, análise qualitativa de dados e um mapeamento sistemático sobre a aplicação de análise qualitativa de dados em revisão sistemática. Além da fundamentação teórica sobre as técnicas de revisão sistemática e análise qualitativa de dados, também são descritas ferramentas que oferecem suporte computacional para aplicação de cada uma das técnicas separadas, entre elas as ferramentas StArt e Insight - utilizadas como base computacional para condução deste trabalho.

No capítulo 3 é descrito a caracterização da análise temática, técnica de análise qualitativa

escolhida como mais adequada à aplicação em revisão sistemática, no contexto de revisão sistemática. É apresentada a descrição da abordagem RS/AT, proposta neste trabalho, bem como a integração e desenvolvimentos realizadas nas ferramentas StArt e Insight.

No capítulo 4 são descritos os estudos realizados a fim de embasar e confirmar as hipóteses obtidas a partir dos estudos conceituais realizados e descritos no capítulo 2.

Finalmente, o capítulo 5 apresenta a conclusão do trabalho, contribuições, limitações e trabalhos futuros a serem realizados nesse tema abordado.

# Capítulo 2

## FUNDAMENTAÇÃO

---

---

### 2.1 Considerações iniciais

Conforme visto no capítulo anterior, este trabalho está inserido no contexto de análise qualitativa de dados em estudos secundários.

Para melhor entendimento do leitor, nas seções abaixo foram descritos os conceitos utilizados na elaboração do trabalho: estudos secundários (revisão sistemática e mapeamento sistemático) e análise qualitativa de dados.

Este capítulo está organizado de acordo com os tópicos a seguir. A seção 2.2 apresenta a fundamentação sobre estudos secundários. Na seção 2.2.1 é detalhado o processo de revisão sistemática. Na seção 2.2.2 é detalhado o processo de mapeamento sistemático. Na seção 2.2.3 são apresentadas ferramentas que podem ser utilizadas para auxiliar a condução de estudos secundários. Na seção 2.2.4 é detalhada a ferramenta StArt, ferramenta que auxilia a condução de revisão sistemática e utilizada neste trabalho. A partir da seção 2.3 são apresentados os conceitos relativos à análise qualitativa de dados. A seção 2.3 apresenta a fundamentação sobre análise qualitativa de dados. Na seção 2.3.1 é abordada a utilização da análise qualitativa na área de Engenharia de Software. Na seção 2.3.2 são apresentadas técnicas para auxiliar a condução da análise qualitativa de dados. Na seção 2.3.3 são apresentadas ferramentas que podem ser utilizadas para auxiliar a análise qualitativa de dados. Na seção 2.3.4 é detalhada a ferramenta Insight, utilizada para auxiliar a análise qualitativa de dados e utilizada em conjunto com a StArt neste trabalho. Finalmente, a seção 2.4 apresenta um mapeamento sistemático conduzido sobre utilização da análise qualitativa de dados em revisões sistemáticas.

## 2.2 Estudos secundários

Para se desenvolver uma teoria ou processo, é necessária coleta de dados, investigação da área de interesse, análise e triangulação das informações entre outros aspectos. A fim de obter-se tais informações, um dos métodos que podem ser utilizado é a realização de estudos secundários.

Por estudos secundários, entende-se a condução de estudos que visem identificar, avaliar e interpretar todos os resultados relevantes a um determinado tópico, fenômeno de interesse ou questão de pesquisa (KITCHENHAM, 2004). São utilizados para sintetizar os estudos primários relativos a um tópico de interesse a fim de obter-se o Estado da Arte daquele tópico.

As etapas simplificadas para condução de um estudo secundário são:

1. *Coleta das evidências*: O primeiro passo é realizar a coleta do material de interesse para o estudo. As evidências, ou estudos primários, podem ser obtidas manualmente ou com auxílio de indexadores de estudos, as chamadas bases de dados eletrônicas (Exemplo: Scopus, ACM, IEEE, entre outras).
2. *Análise dos dados*: Após coletados os estudos primários, estes devem ser analisados baseado em critérios definidos pelo pesquisador para obtenção de informações relevantes ao escopo tratado.
3. *Obtenção do Estado da Arte*: Por fim, deve-se sumarizar todas as informações coletadas e obter o Estado da Arte do tema proposto no estudo secundário.

As mencionadas etapas para condução de um estudo secundário são comuns a qualquer estudo deste tipo. Nesta categoria de estudos, destacam-se a Revisão Sistemática e Mapeamento Sistemático, descritos na sequência.

### 2.2.1 Revisão Sistemática

A revisão sistemática é um tipo de estudo secundário realizado por meio de passos sistematicamente definidos por Kitchenham (2004). Como é um estudo secundário, seu objetivo é sumarizar os estudos primários relativos ao escopo de interesse definido.

Na revisão tradicional, os métodos utilizados variam de acordo com a experiência e habilidade de cada pesquisador. Já a revisão sistemática auxilia a documentação e condução do

processo por meio dos passos definidos, tornando possível que o mesmo seja replicado por outros pesquisadores.

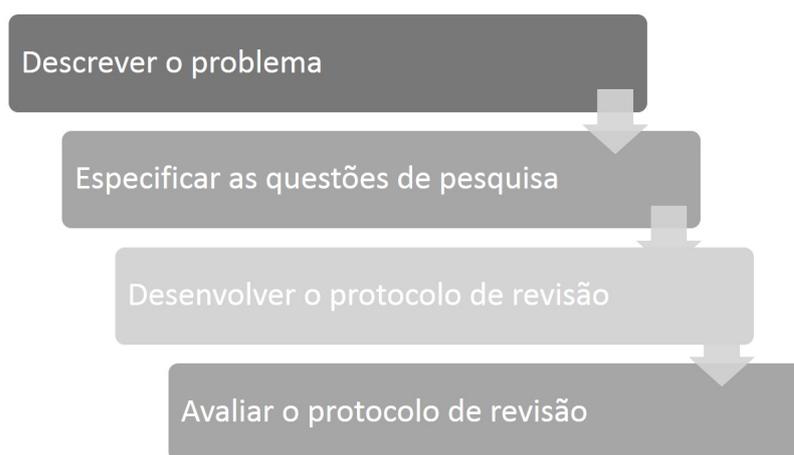
Kitchenham (2004) define a condução de uma revisão sistemática nas seguintes etapas e suas atividades:

1. *Planejamento* - Neste passo devem ser realizadas as seguintes etapas:

- **Identificação da necessidade de uma revisão sistemática:** Efetuar uma revisão sistemática não é uma tarefa trivial, leva tempo e dedicação de quem a conduz. Portanto, é necessário definir o que se deseja obter na mesma e a real necessidade de uma extensa busca na literatura.
- **Desenvolvimento de um protocolo de revisão:** No protocolo da revisão se estabelece o motivo da revisão sistemática, as fontes dos dados a serem buscados, os parâmetros de seleção e aceitação dos estudos, formulários de extração e sumarização dos dados.

A medida que a revisão é conduzida, o protocolo pode ser alterado a fim de se adequar melhor aos critérios e ao processo de busca dos dados, porém, os dados previamente obtidos devem ser revisados para não haver inconsistências na pesquisa.

O processo simplificado de desenvolvimento do protocolo de revisão está ilustrado na Figura 2.1.



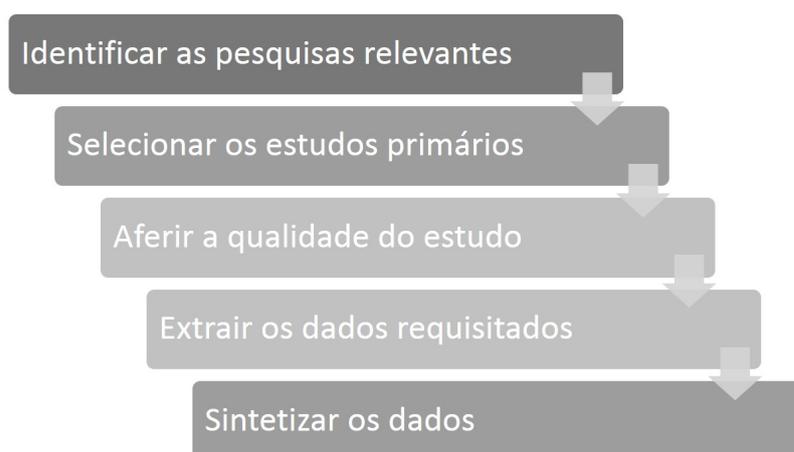
**Figura 2.1: Desenvolvimento do protocolo da Revisão Sistemática**

2. *Condução* - Este é o passo no qual o processo de busca, análise, extração e sumarização dos dados será executado. O fluxo básico da condução da revisão sistemática é ilustrado na Figura 2.2. Cada uma das etapas envolvidas são detalhadas a seguir:

- **Identificação da pesquisa:** Consiste em detalhar as fontes das quais foram obtidos os estudos a serem analisados bem como a composição das palavras e condições de busca utilizadas nas bases eletrônicas de dados (*string* de busca). Esta etapa possui forte ligação com o fato da revisão sistemática ser um processo que pode ser repetido por outro pesquisador, portanto, é necessária toda a documentação dos processos utilizados para condução da mesma.
- **Seleção dos estudos:** No protocolo definido durante a etapa de planejamento foram definidos critérios para aceitação ou exclusão dos estudos, estes critérios serão utilizados nesta etapa. Por meio da leitura do resumo dos artigos e palavras-chave, é efetuada uma pré-seleção. Nesta primeira etapa de seleção os artigos que não atenderem aos critérios estabelecidos serão rejeitados.
- **Avaliação da qualidade dos estudos:** Após a primeira etapa de seleção é conduzida outra seleção. Os materiais que chegaram a esta etapa serão novamente avaliados a fim de identificar a relevância das publicações para o estudo que está sendo conduzido bem como a credibilidade do material aceito. Nesta, o estudo é lido integralmente e marcado como aceito ou rejeitado.  
A avaliação da qualidade dos estudos deve ter rigor, credibilidade e relevância (DYBA; DINGSOYR; HANSEN, 2007).
- **Extração dos dados:** Na etapa de extração de dados, geralmente são utilizados formulários pré-estabelecidos pelos pesquisadores a fim de guiar a coleta dos dados nos estudos e evitar que dados relevantes para a pesquisa passem despercebidos por quem conduz a revisão.  
Caso a revisão seja conduzida por mais de um revisor, os formulários servem de parâmetro para dados a serem extraídos, evitando assim discrepâncias dos dados coletados.
- **Síntese dos dados:** Após todo o processo sistemático e de posse dos dados extraídos, os dados devem ser sintetizados para obter o estado da arte do tópico abordado na revisão.

3. *Documentação* - Após concluída a revisão sistemática proposta, o processo e os resultados obtidos devem ser publicados em meio adequado (relatório técnico, artigo de periódico ou conferência, entre outros) para transmitir o conhecimento adquirido.

A aplicação das mencionadas etapas caracteriza a realização de uma revisão sistemática. É um processo árduo e trabalhoso, porém, traz credibilidade à pesquisa visto o rigor metodológico com o qual é conduzida a síntese dos estudos primários.



**Figura 2.2: Etapas da condução da Revisão Sistemática**

Além dos trabalhos de Kitchenham (2004), os trabalhos de Fabbri et al. (2013) também abordam a definição de etapas e procedimentos para realização de uma revisão sistemática. Neste, o processo de revisão sistemática estabelecido é um processo iterativo, chamado de revisão sistemática incremental (I-SR), no qual a realização das etapas de planejamento, seleção inicial, seleção final e extração, e síntese. Consiste da aplicação de ciclos de atividades a serem realizadas, conforme ilustrado na Figura 2.3.

### 2.2.2 Mapeamento sistemático

De acordo com Petticrew e Roberts (2008), o mapeamento sistemático é um método projetado para fornecer uma visão ampla de determinada área de pesquisa, que permite identificar, quantificar e analisar os resultados, estabelecendo evidências da mesma sobre determinado tema. É uma pesquisa de literatura que visa identificar quais tipos de estudos podem ser abordados por uma Revisão Sistemática, indicar o local no qual os mesmos foram publicados, em que bases de dados foram indexados e quais os tipos de resultados apresentaram.

Fornecer 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 (PETERSEN et al., 2008).

É confundido algumas vezes com revisão sistemática. O mapeamento sistemático também é um estudo secundário e é similar a revisão sistemática. Porém, a principal característica que diferencia um mapeamento sistemático de uma revisão sistemática é o enfoque com o qual é conduzido. O mapeamento sistemático realiza uma revisão mais ampla dos estudos primários, em busca de identificar quais evidências estão disponíveis, bem como identificar lacunas no

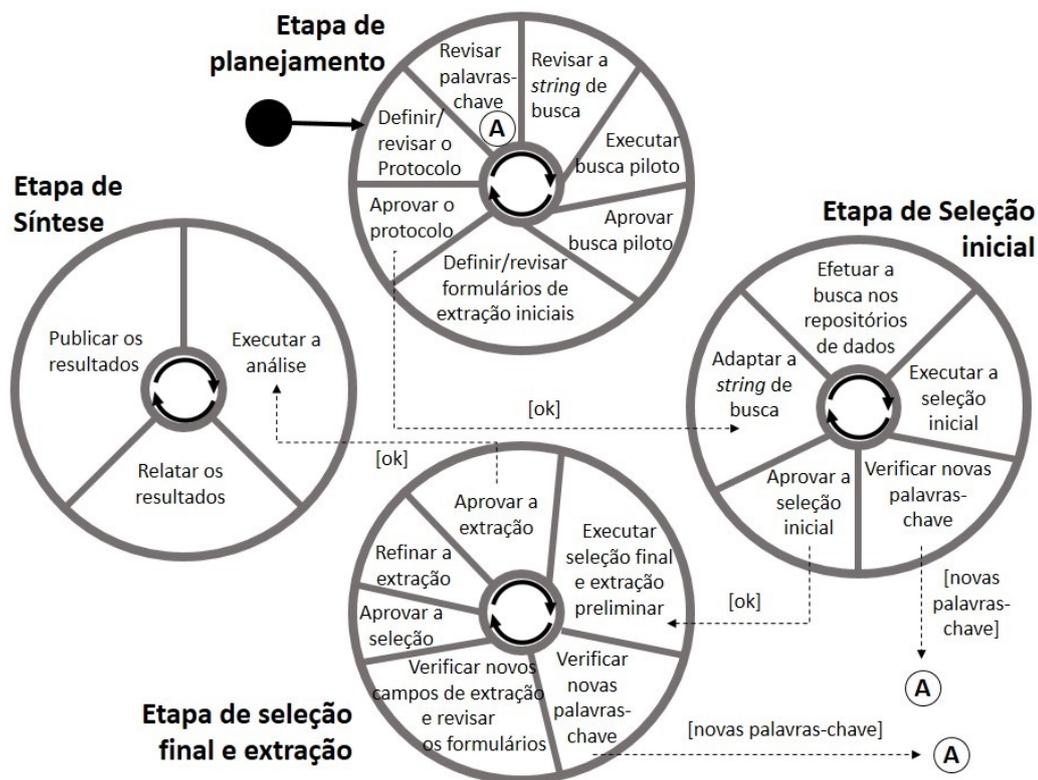


Figura 2.3: Etapas da Revisão Sistemática Iterativa - adaptado de Fabbri et al. (2013)

conjunto dos estudos primários. A realização de um mapeamento pode direcionar o foco de revisões sistemáticas futuras e identificar áreas nas quais há necessidade de se conduzir outros estudos primários (KITCHENHAM, 2004).

A principal diferença entre os métodos é o resultado a ser obtido: na revisão sistemática a obtenção do estado da arte sobre um tema enquanto que no mapeamento sistemático a busca é pelo panorama geral da área de estudo.

### 2.2.3 Ferramentas para auxiliar a condução de Estudos secundários

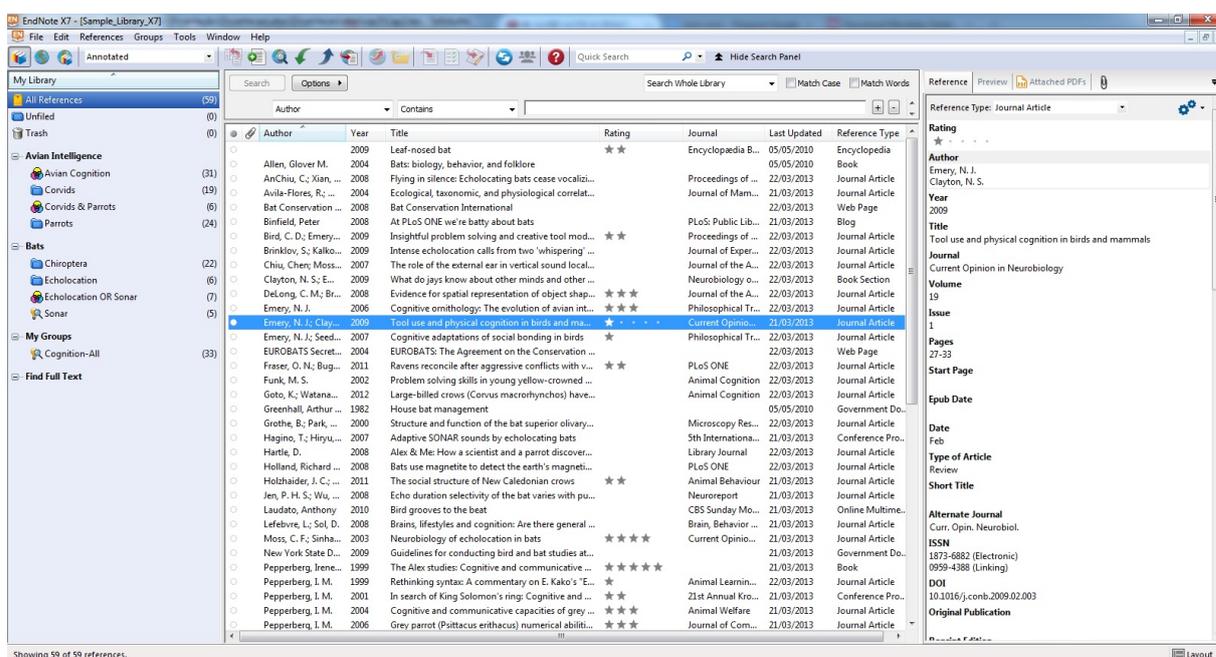
Conduzir um estudo secundário é uma tarefa que exige rigor metodológico e documentação do processo. Para auxiliar a condução de tal, o uso de gerenciadores de referências e, principalmente, ferramentas que auxiliem todas as etapas do processo, são importantes recursos para esta tarefa.

No quesito de gerenciadores de referências, merecem destaque os seguintes:

- *Endnote*: Como mencionado, a ferramenta EndNote (Figura 2.4) não é em si uma ferramenta específica para condução de estudos secundários mas sim um gerenciador de

referências. Foi desenvolvida pela Thomson Reuters e não é gratuita. É usada para auxiliar o processo da revisão sistemática pois permite organizar, manipular e compartilhar com outros pesquisadores as referências coletadas. Possui um plugin disponível para o MS Word que auxilia a organização das referências utilizadas.

- **Mendeley:** É um gerenciador de referência gratuito (Figura 2.5). Possui opções de importar documentos locais, buscar online outros e realizar anotações nos documentos. Assim como o Endnote, possui um plugin para o MS Word.
- **Zotero:** Suas funcionalidades são similares às dos mencionados gerenciadores de referência. Seu funcionamento pode ser realizado via Web por meio da instalação de um plugin no navegador Mozilla Firefox ou como um aplicativo separado que se conecta aos navegadores Google Chrome, Safari ou mesmo o Mozilla Firefox.



**Figura 2.4: Gerenciador de referências EndNote**

Quanto a ferramentas que auxiliem nas etapas dos estudos secundários, existem alternativas gratuitas e comerciais para esta tarefa. Comercialmente, uma das ferramentas mais utilizadas é a EPPI Reviewer, desenvolvida e comercializada pelo Instituto de Educação da Universidade de Londres. Em relação a ferramentas gratuitas, destacamos a Revis, desenvolvida pelo Instituto de Ciências Matemáticas e de Computação da USP, e a StArt, desenvolvida pelo Laboratório de Pesquisa em Engenharia de Software da UFSCar. Uma breve descrição de cada uma das mencionadas é apresentada a seguir:

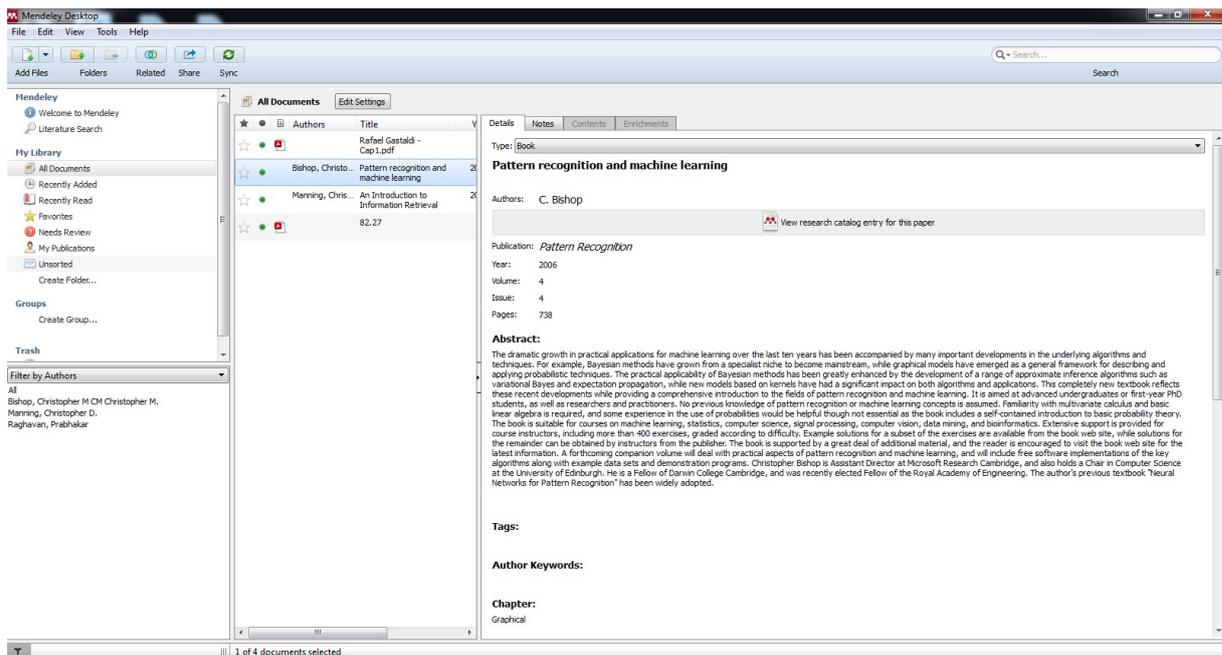


Figura 2.5: Gerenciador de referências Mendeley

- *Eppi Reviewer*: Esta ferramenta é um desenvolvimento do Centro de Coordenação de evidência para a política e informação (*EPPI-Centre*), parte da Unidade de Pesquisa em Ciências Sociais no Instituto de Educação da Universidade de Londres. É uma ferramenta online baseada na tecnologia Silverlight da Microsoft. Dá suporte a seleção, extração e sumarização de estudos secundários. A *Eppi Reviewer* (Figura 2.6) possui uma versão para uso limitado gratuito de 30 dias. Após o período é necessário aquisição da ferramenta para continuar a utilizá-la.

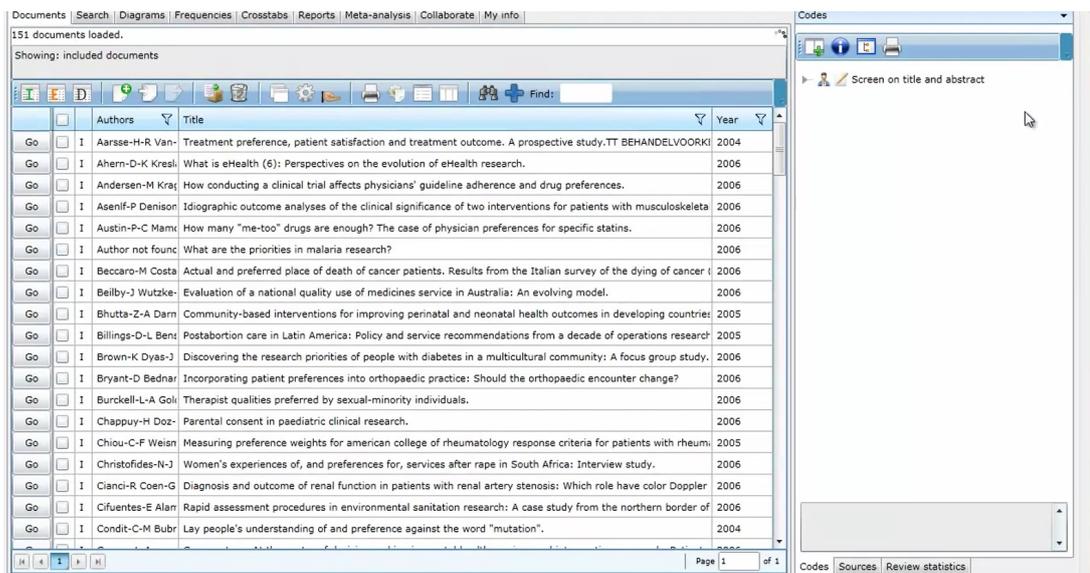


Figura 2.6: Ferramenta Eppi Reviewer

- **Revis:** A ferramenta Revis (Figura 2.7) foi desenvolvida pelo Instituto de Ciências Matemáticas e de Computação da USP. É baseada em técnicas de visualização e mineração de texto, denominado *Visual Text Mining* (FELIZARDO et al., 2011). Entre as técnicas de visualização implementadas nesta ferramenta, destacam-se a rede de citações (*citation network*) e *edge bundles*. Auxiliam de forma visual a verificar as relações entre os documentos e suas citações. É gratuita e dá suporte, principalmente, à fase de seleção dos estudos com o auxílio das técnicas de *Visual Text Mining*.

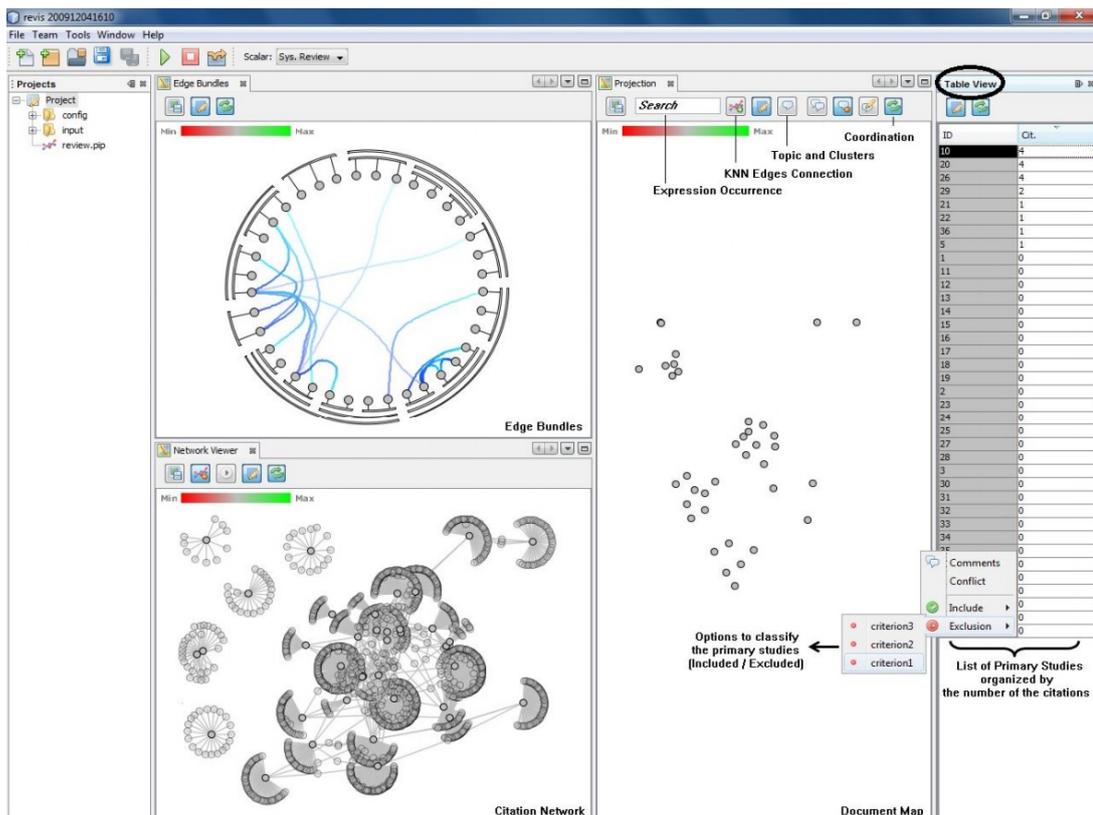


Figura 2.7: Ferramenta Revis - retirado de Felizardo et al. (2011)

- **StArt:** A ferramenta StArt foi desenvolvida pelo Laboratório de Pesquisa em Engenharia de Software da UFSCar para auxiliar o processo de revisão sistemática. Esta ferramenta foi utilizada neste trabalho e será detalhada na seção a seguir.

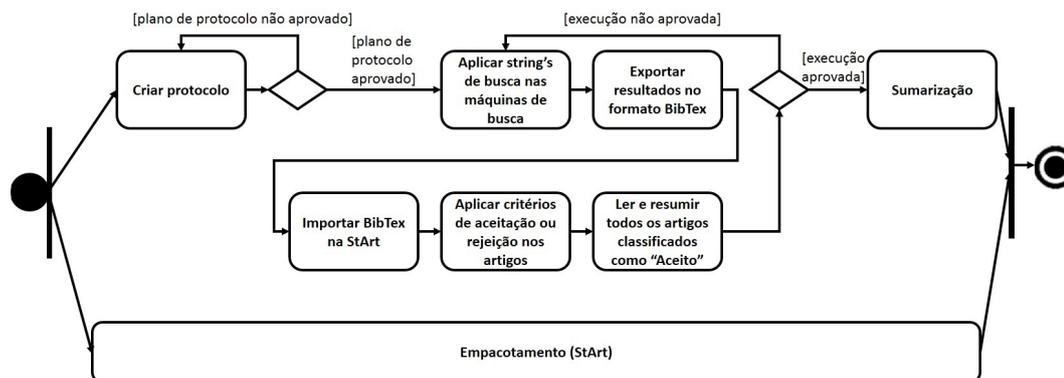
As citadas ferramentas e outras não listadas, seja gerenciadores de referência ou específicas para estudos secundários, são importantes para auxiliar a condução de um estudo secundário. O volume de material adquirido é grande e gerenciar manualmente todos os dados é uma tarefa difícil. Além disso auxiliam a documentação do estudo, requisito obrigatório para um estudo sistemático visto que deve fornecer condições de ser replicado por outro pesquisador.

Um estudo comparando diversos tipos de ferramentas com o propósito de auxiliar a condu-

ção da revisão sistemática, entre elas a ferramenta StArt, foi realizado por Marshall, Brereton e Kitchenham (2014). Neste estudo são elencadas algumas vantagens e desvantagens da utilização de cada ferramenta para auxiliar a condução de revisões sistemáticas.

### 2.2.4 StArt

Conforme mencionado, a ferramenta StArt foi desenvolvida para auxiliar o processo de revisão sistemática. Possibilita ao revisor, do menos ao mais experiente, um processo guiado e documentado do trabalho. O processo para revisão sistemática utilizado na ferramenta segue os passos definidos por Kitchenham (2004) e é descrito por Zamboni et al. (2010), conforme a Figura 2.8.



**Figura 2.8: Processo simplificado para utilização da StArt, adaptado de Zamboni et al. (2010)**

Vale ressaltar que, apesar de uma sequencia de etapas ser estabelecida neste fluxo, o processo de revisão sistemática definido, assim como a ferramenta, possibilita um processo iterativo no qual ocorre a navegação entre as etapas (conforme descrito no Capítulo 2, destacando a Revisão Sistemática Incremental - I-SR).

A organização inicial ferramenta é apresentada na Figura 2.9. O trecho destacado com borda amarela e rotulado com a letra A destaca o guia de navegação da ferramenta. Nesta aba são mostradas as informações relativas às etapas da revisão e a quantidade de artigos aceitos, rejeitados ou duplicados em cada etapa do processo. O trecho destacado com borda vermelha e rotulado com a letra B apresenta a etapa corrente da revisão - são destacadas as informações específicas da etapa que está sendo conduzida.

Ao iniciar a revisão sistemática com a ferramenta, o primeiro passo é a criação do projeto da revisão, na qual são preenchidas as informações básicas: pesquisadores envolvidos, título

da revisão e descrição do trabalho a ser conduzido. Após criado o projeto é gerado um arquivo com extensão .start, no qual todas as informações da revisão são gravadas e é possível o compartilhamento da mesma (basta enviar o arquivo ao interessado para abrir na ferramenta StArt). Criado o projeto, dará-se início ao planejamento com o preenchimento do protocolo da revisão. No protocolo serão estabelecidos o objetivo principal, máquinas de busca escolhidas, palavras-chave utilizadas para composição das *strings* de busca, critérios de inclusão e exclusão dos estudos, bem como as demais informações relativas à condução da revisão.

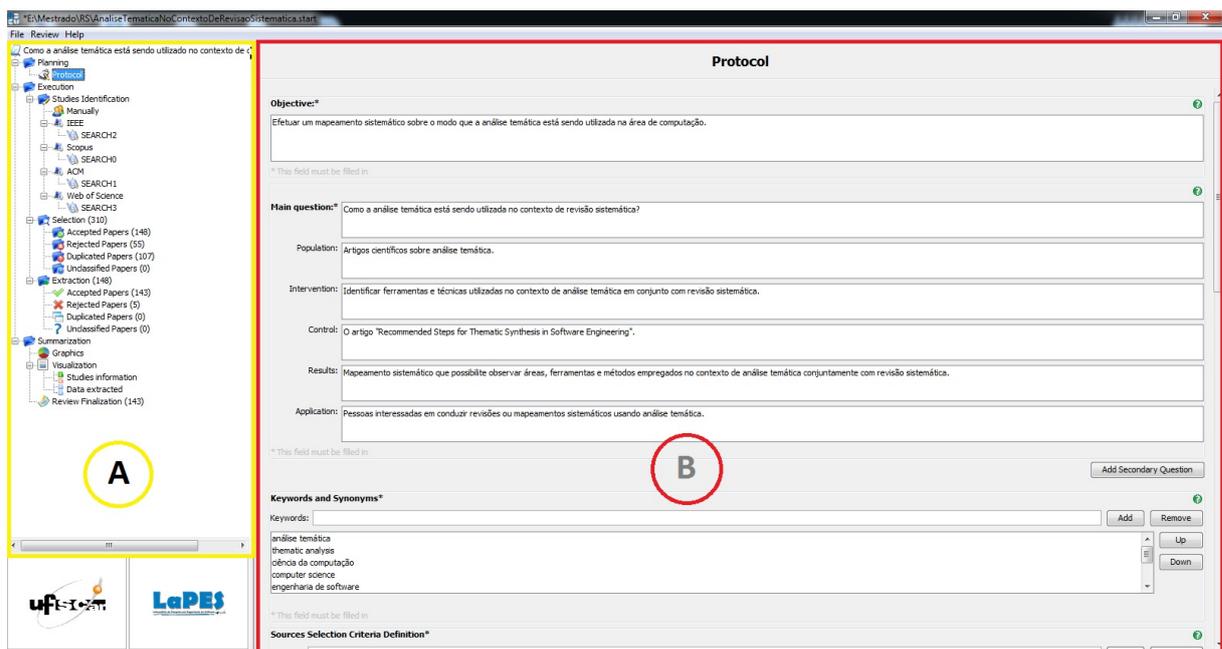


Figura 2.9: Organização da ferramenta StArt e definição do protocolo

Definido o protocolo da revisão, inicia-se a importação dos estudos primários. Pode ser realizado de dois modos: importar os estudos na ferramenta manualmente ou por meio da criação de uma seção de busca relativa às bases indexadoras de artigo (IEEE, ACM, Scopus, etc).

No caso da importação manual de um estudo, cabe ao pesquisador preencher os dados do estudo primário em questão, como resumo, título, pesquisadores envolvidos e outros campos solicitados pela ferramenta. Já no caso da importação de dados advindos de uma base indexadora, o pesquisador deverá criar uma nova seção de busca e preencher os dados relativos à *string* de busca utilizada e base indexadora na qual foi aplicada, conforme a Figura 2.10. Após preenchido os dados da base indexadora, deve-se importar o arquivo de extensão .bibtex (exportado pelas bases indexadoras) na seção de busca relativa à base. O arquivo .bibtex contém as informações relativas aos estudos primários (autores, título, local de publicação, resumo, entre outros) e é interpretado automaticamente pela ferramenta.

Uma vez importado os estudos primários na ferramenta, estes serão listados nesta etapa de

Seleção com o status definido como "Unclassified", ou não classificado. Será então iniciado o processo de seleção dos estudos. Ao clicar duas vezes sobre um estudo listado, será mostrada uma janela, conforme mostrado na Figura 2.11, contendo as informações relativas ao estudo. O pesquisador deverá então aplicar os critérios de inclusão e exclusão definidos para cada estudo e classificá-los como "Aceito" ou "Rejeitado".

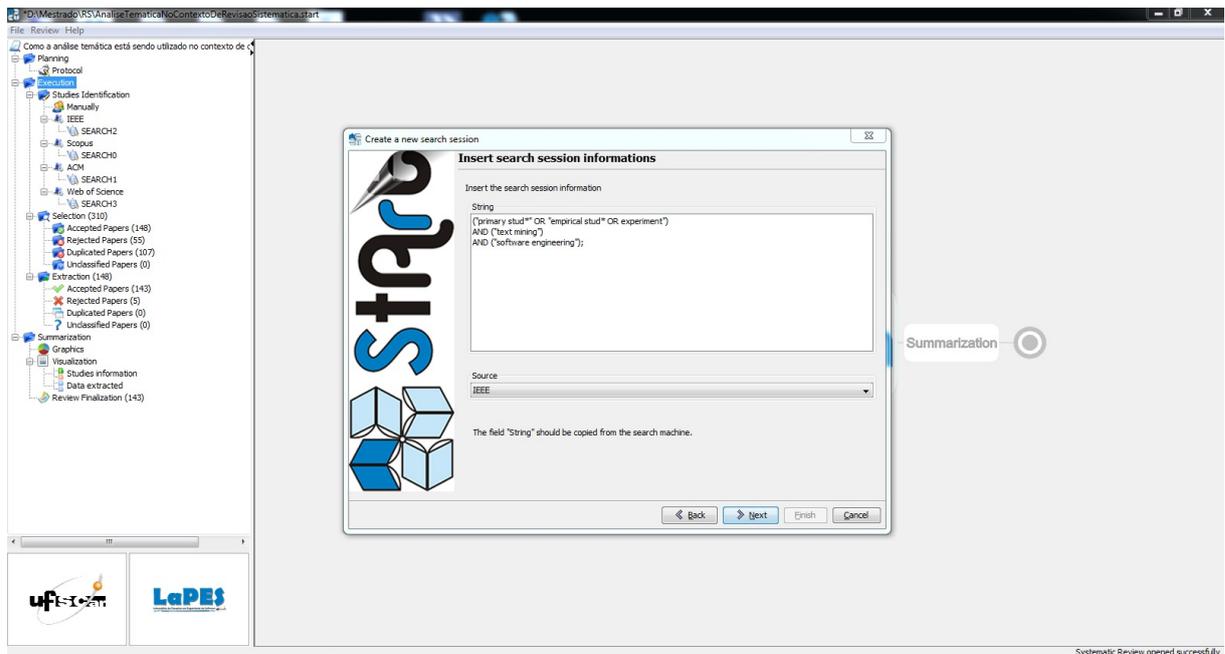


Figura 2.10: Adicionar nova seção de busca

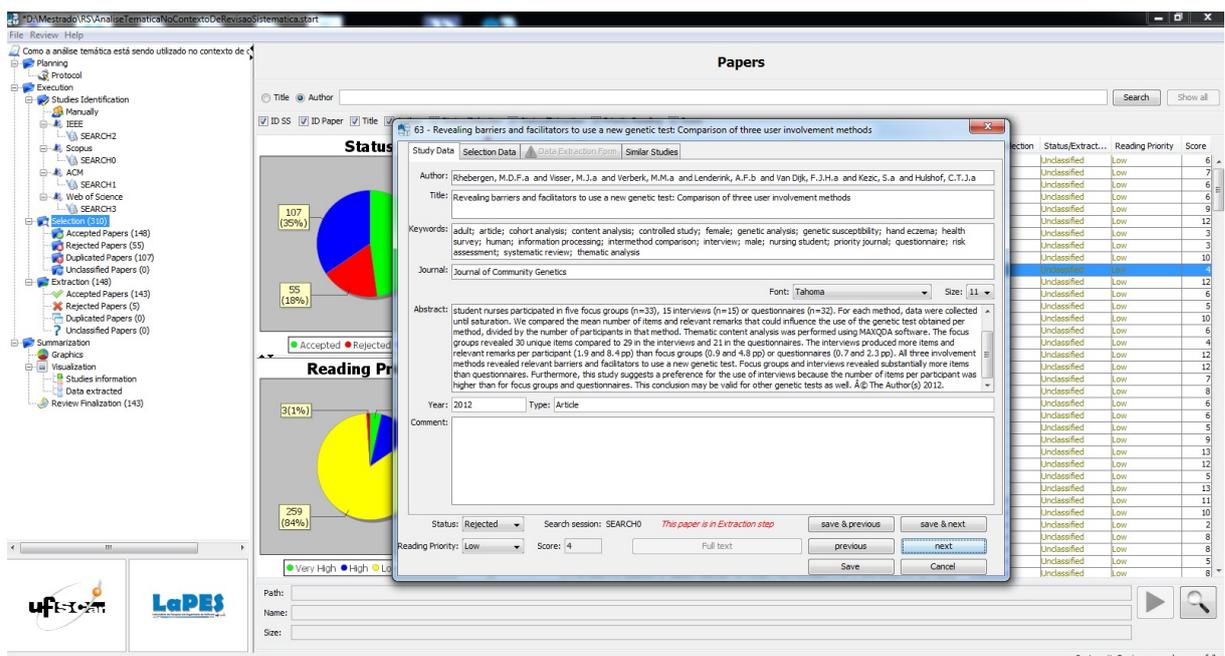


Figura 2.11: Etapa de seleção dos estudos

Selecionados os estudos, os que foram marcados como "Aceito" na etapa de Seleção, são

listados nesta etapa de Extração. Nesta etapa, o autor deve adquirir o estudo completo (*full text*), efetuar a leitura completa do mesmo e executar a extração dos dados.

No processo tradicional da revisão sistemática, esta extração é realizada por meio dos formulários de extração de dados, definidos no protocolo e em destaque na Figura 2.12. Esta é uma das etapas da revisão que pode gerar maior quantidade de retrabalho pois, a medida que os estudos são lidos na íntegra, novos campos de extração podem ser identificados. Caso ocorra a mencionada situação, os estudos analisados anteriores deverão ser relidos a fim de evitar inconsistências nos dados extraídos.

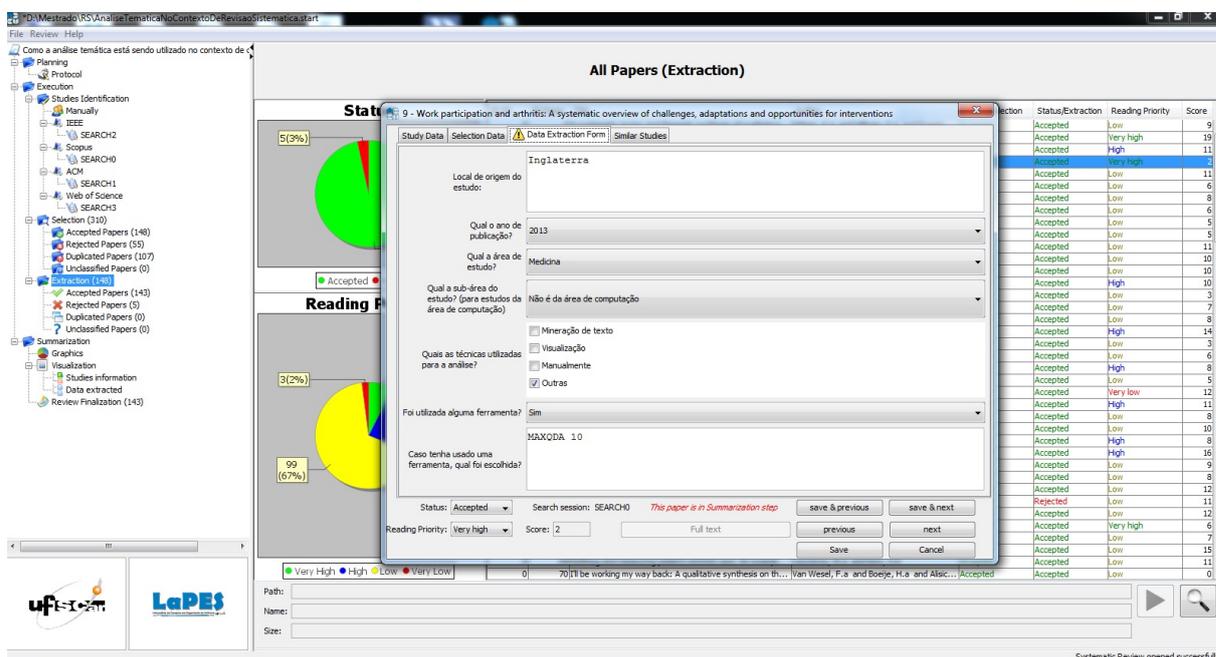


Figura 2.12: Formulários de extração na StArt

Por fim, como última etapa a ser conduzida executa-se a sumarização dos dados. Para tal, todos os dados definidos para a extração devem ter sido coletados. A ferramenta StArt disponibiliza a geração de relatórios com os dados extraídos e informações relativas ao artigo bem como diferentes modelos de visualização para os dados coletados (Figura 2.13).

## 2.3 Análise qualitativa

Strauss e Corbin (2008) definem o termo "pesquisa qualitativa" como qualquer tipo de pesquisa que produza resultados não alcançados por meio de procedimentos estatísticos ou de outros meios de quantificação. Pode se referir à pesquisa sobre a vida das pessoas, experiências vividas, comportamentos, emoções e sentimentos, e também à pesquisa sobre o funcionamento organizacional, movimentos sociais, fenômenos culturais e interação entre nações. Alguns da-

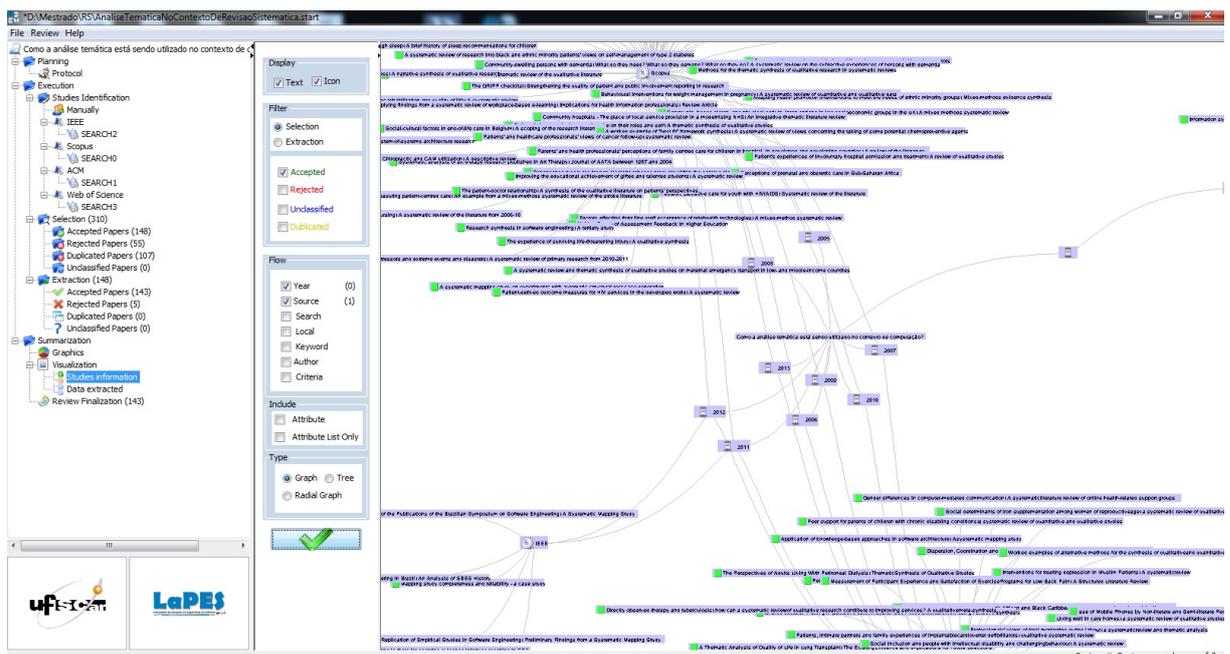


Figura 2.13: Visualização de dados para sumarização

dos podem ser quantificados, como no caso do censo ou de informações históricas sobre pessoas ou objetos estudados, mas o principal da análise é interpretativa.

A pesquisa qualitativa tem por objetivo traduzir e expressar o sentido dos fenômenos do mundo social, trata-se de reduzir a distância entre indicador e indicado, entre teoria e dados, entre contexto e ação (MAANEN, 1979). Consistem em analisar e correlacionar dados a fim de obter-se informações não quantificáveis.

É muito empregada em ciências sociais, medicina e vem ganhando espaço nos estudos de engenharia de software por ser mais adequada quando aplicada em resultados não quantificáveis levando em consideração fatores externos (população, modo de coleta dos dados, contexto, entre outros) e interpretação dos resultados obtidos. Orlikowski e Baroudi (1991) citam que desde 1991 pesquisadores em sistemas da informação, interação humano-computador e engenharia de software começavam a utilizar os métodos qualitativos.

Visto que grande parte da análise qualitativa busca correlacionar dados não quantificáveis, o processo de análise qualitativa dos dados não é algo trivial. Os materiais são dos formatos mais variados, desde fragmentos de texto e textos completos até entrevistas ou vídeos pertinentes ao assunto.

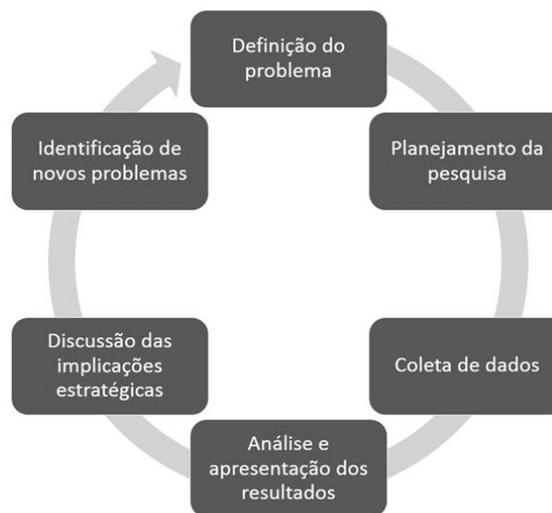
De acordo com Seaman (1999), dois conjuntos de métodos de análise de dados qualitativos podem ser identificados:

- **Geração de teoria:** Método usado para gerar hipóteses baseadas em dados. Por exemplo,

método de comparação constante e análise cruzada.

- **Confirmação de teoria:** Método usado para construir "evidências de peso" necessárias para confirmar a hipótese. O objetivo desses métodos não são provar a teoria. Por exemplo, validação, triangulação, anomalias nos dados, análise de caso negativo e replicação.

O processo de análise qualitativa exige muito trabalho e dedicação. O material de estudo pode ser composto por imagens, vídeos, áudios e, principalmente, arquivos textuais. Visto a heterogeneidade do material, realizar a análise não é uma tarefa trivial. É um processo cíclico (Figura 2.14) que demanda planejamento, coleta e interpretação de dados e a geração de temas/problemas obtidos da pesquisa.



**Figura 2.14:** Ciclo da pesquisa qualitativa

A opção pela abordagem qualitativa está diretamente relacionada a quais informações buscase obter da pesquisa a ser realizada. Entre razões para se realizar a opção pelo método qualitativo, Strauss e Corbin (2008) citam várias razões válidas para fazer tal. Uma delas é preferência e/ou experiência dos pesquisadores. Algumas pessoas são mais orientadas e têm temperamento mais apropriado para fazer esse tipo de trabalho. Outra razão, provavelmente mais válida, é a natureza do problema de pesquisa. Por exemplo, uma pesquisa que tenta entender o significado ou a natureza de experiências com pessoas - para tal deve-se levar em conta variáveis não quantificáveis como local, experiência prévia, entre outras. Além disso, os métodos qualitativos podem ser usados para obter detalhes intrincados sobre fenômenos como sentimentos, processos de pensamento e emoções que são difíceis de extrair ou de descobrir por meio de métodos de pesquisa mais convencionais.

O aumento no uso de análise qualitativa tem se intensificado na área de Engenharia de

Software visto que o comportamento humano pode influenciar no uso das técnicas da área e por se tratar de uma área que envolve um misto de aspectos técnicos e humanos (SEAMAN, 2008). A subseção a seguir detalha brevemente um panorama da técnica na área em questão.

### 2.3.1 Análise qualitativa em Engenharia de Software

De acordo com Travassos (2002), quando falamos de Engenharia de Software, a discussão concentra-se em se podemos considerá-la ciência ou engenharia. Esta questão refere ao duplo caráter do software. Por um lado, a Engenharia de Software considera o processo de criação do produto (software) e, desse ponto de vista, ela possui as características explícitas de produção ou engenharia. Por outro lado, os aspectos relacionados a *time-to-market* e competição exigem a melhoria contínua e sequencial da qualidade do processo e do produto. É nesse contexto que a parte científica de Engenharia de Software se apresenta. Em vista disso, para atingir um nível adequado de evidência a respeito da caracterização de uma determinada tecnologia em uso, a Engenharia de Software Baseada em Evidências deve fazer uso basicamente de dois tipos de estudos: estudos primários e estudos secundários (MAFRA; TRAVASSOS, 2006).

Para se efetuar a análise dos citados estudos secundários, técnicas para aumentar a produtividade e qualidade dessas análises são fundamentais. Como mencionado anteriormente neste trabalho, o uso de análise qualitativa de dados tem sido mais utilizado pelas áreas de ciências humanas e da saúde. Porém, seu uso pode trazer relevantes reflexões e resultados a serem obtidos dos materiais de estudo. Myers (1997) cita que, em Sistemas de Informação, tem havido uma guinada do enfoque de pesquisa de questões puramente tecnológicas para questões gerenciais e organizacionais, fazendo com que o interesse pela aplicação de métodos qualitativos seja incrementado.

Kitchenham (2004) ressalta que a Engenharia de Software Baseada em Evidência deve prover meios pelos quais a melhores evidências provenientes da pesquisa possam ser integradas com experiência prática e valores humanos no processo de tomada de decisão considerando o desenvolvimento e a manutenção do software. Como resultado, Travassos (2002) cita os seguintes pontos que encorajam o desenvolvimento de técnicas para auxiliar a análise de dados:

- Organizações de desenvolvimento de software seriam contempladas com um embasamento científico que apoiasse a tomada de decisão no que se refere à aquisição de novas tecnologias alinhadas às suas necessidades. Desse modo, engenheiros de software poderiam ter à sua disposição tecnologias que contribuíssem para o aumento de produtividade, resultando em eventuais diminuições de custos de desenvolvimento e melhorias de quali-

dade nos produtos.

- Usuários de software poderiam ter à sua disposição produtos de alta qualidade, desenvolvidos com tecnologias adequadas, que atendessem às suas necessidades. De modo similar, o público em geral poderia ser beneficiado com o funcionamento de produtos confiáveis que não representassem eventuais riscos à sua segurança e bem-estar, nem desperdício de dinheiro público por parte dos governantes.

O aumento no uso de análise qualitativa tem se intensificado na área de Engenharia de Software visto que o comportamento humano pode influenciar no uso das técnicas da área e por se tratar de uma área que envolve um misto de aspectos técnicos e humanos (SEAMAN, 2008). Ressalta-se também que esse tipo de análise traz relevantes informações para pesquisadores da área. Por exemplo, ao se comparar técnicas conhecer os motivos de "por quê?" uma é mais efetiva que a outra e não só "qual?" é a mais efetiva.

Alguns pesquisadores consideram o tratamento do "contexto" tão importante quanto o aspecto técnico do objeto de estudo. Tais fatores são necessários serem avaliados pois software não existe no vazio. Kaplan e Maxwell (2005) citam que sua implementação, uso e sucesso ou fracasso ocorrem em um contexto social e organizacional que molda o que acontece quando esse sistema é introduzido.

Para auxiliar a condução da análise existem softwares específicos para este propósito. Conforme mencionado anteriormente, a análise qualitativa não é um processo trivial, portanto, técnicas computacionais que possam contribuir para esta tarefa são importantes. Nas subseções a seguir são descritas algumas técnicas e ferramentas utilizadas para análise qualitativa.

### **2.3.2 Técnicas para auxiliar a condução da análise qualitativa**

Como já citado anteriormente, aplicar análise qualitativa não é um processo simples e está fortemente associado à habilidade do pesquisador que conduz o estudo para obtenção de resultados satisfatórios. Com objetivo de otimizar e guiar o processo de análise, várias técnicas foram e estão sendo desenvolvidas sobre o assunto.

Vale ressaltar que muitas delas podem ser utilizadas em conjunto - como exemplo é possível citar a técnica de *coding*, uma das técnicas mais utilizadas em conjunto com as outras pois auxilia o pesquisador na extração e categorização dos fragmentos dos textos em categorias. Nos tópicos abaixo serão descritos os principais pontos de algumas delas.

- **Coding**

Independente do método de análise qualitativo empregado, a técnica de *coding*, ou codificação dos documentos, apoia o pesquisador na análise dos dados.

A ideia principal é atribuir rótulos a passagens relevantes dos dados para permitir organizá-los e correlacioná-los em temas e categorias. Seaman (2008) divide a aplicação da técnica em 3 etapas:

- *Open coding*: O primeiro passo na construção da teoria é a conceituação - representação abstrata de um fato, objetivo ou ação/interação que um pesquisador identifica como importante nos dados. Embora fatos ou acontecimentos possam ser elementos distintos, o fato de compartilharem características comuns ou significados relacionados permite que sejam agrupados.

Nesta etapa o pesquisador deve ler o material buscando referências a respeito do tópico de interesse e inserir rótulos (códigos) para cada trecho extraído relevante (*quotation*).

- *Axial coding*: O objetivo da codificação axial é começar o processo de reagrupamento dos dados que foram divididos durante a codificação aberta.

Nesta etapa os códigos extraídos devem ser agrupados para gerar categorias que promovam melhor entendimento dos dados. Embora a codificação axial tenha objetivo diferente da codificação aberta, esses passos não são necessariamente sequenciais - o sentido que as categorias se relacionam surge a medida que se efetua a codificação aberta.

- *Selective coding*: Por fim, o pesquisador deve analisar novamente os códigos e categorias obtidos e elaborar uma descrição que sintetize os dados analisados.

Existem diversas técnicas para efetuar esse correlacionamento. O primeiro e fundamental passo é identificar a categoria central - também chamada de categoria básica, representa o tema principal da pesquisa e todas as outras categorias importantes podem ser relacionadas a ela.

As etapas descritas acima tratam de modo mais abrangente a aplicação da técnica. Hancock, Ockleford e Windridge (1998) definem de modo mais detalhado a aplicação da técnica por meio dos seguintes passos:

1. Ler os dados em forma textual procurando trechos (citações) com informações relevantes e atribuir um pequeno texto (código) que representa aquele tópico;

2. Elaborar uma lista com todos os códigos diferentes;
3. Agrupar os códigos em categorias que devem representar o principal tópico relacionado aos códigos e elaborar uma lista com essas categorias;
4. Se há categorias relacionadas, criar outra categoria e definir uma hierarquia entre elas;
5. Analisar e comparar todas as categorias, alterando as posições na hierarquia e criando novas categorias se necessário;
6. Repetir os passos 1 a 5 para todos os documentos da pesquisa;
7. Assegurar que os trechos com o mesmo código são relacionados entre si;
8. Assegurar que as categorias, hierarquias e rótulos são representativos;
9. Analisar possíveis relacionamentos entre categorias pois podem surgir importantes reflexões sobre a pesquisa. Essa análise deve ser executada após confirmar que todos os códigos e categorias estão adequados;
10. Revisar os documentos considerando as categorias e procurar trechos não considerados antes mas que agora parecem relevantes.

A aplicação dos passos descritos exige dedicação e cuidado para evitar que dados relevantes passem despercebidos.

Alguns pesquisadores utilizam guias de extração de dados para auxiliar a padronização de algumas categorias e garantir a coleta de dados considerados relevantes ao contexto da pesquisa. Uma das guias que podem ser utilizadas são as *seed categories*, ou categorias sementes na tradução literal - são categorias desenvolvidas inicialmente para auxiliar a categorização dos dados, caso ocorram nos estudos. (BADREDDIN, 2013) cita que a utilização destas categorias semente é um recurso necessário para auxiliar o pesquisador na coleta dos dados.

A utilização da técnica, tal como o da análise qualitativa, demanda dedicação. Consiste de total imersão no material de estudo e constante comparação entre os novos dados com os dados previamente obtidos em um processo iterativo para evitar distorções de resultado.

- **Teoria fundamentada em dados**

A Teoria Fundamentada ou *Grounded Theory* surgiu na década de 1960 desenvolvida por Glaser e Strauss por meio de uma pesquisa sociológica efetuada por eles para análise de mortes em hospitais.

Strauss e Corbin (2008) citam que teorias fundamentadas, por serem baseadas em dados, tendem a oferecer mais discernimento, melhorar o entendimento e fornecer um guia importante para a ação. Tende a se parecer mais com a "realidade" do que a teoria derivada da reunião de uma série de conceitos baseados em experiência ou somente por meio da especulação (como alguém acha que as coisas devem funcionar).

Através de seu trabalho, Glaser e Strauss criaram um método que permitiu ao pesquisador gerar sistematicamente uma teoria substantiva baseada em dados empíricos. O objetivo era desenvolver uma teoria que fosse consistente, adequasse os dados e serviria ao mundo real (WALKER; MYRICK, 2006).

Utilizada principalmente na área da saúde, a técnica baseia-se em, por meio da coleta de dados de diferentes modos (entrevistas, observação, dados estatísticos entre outros), utilizar comparação, codificação e categorização dos mesmos para identificar os relacionamentos contidos.

Um ponto importante que deve ser ressaltado sobre *grounded theory* é que, diferentemente da análise temática, não deve-se usar nenhum guia para extração dos dados. A análise temática sugere que alguns guias podem ser utilizados para auxiliar a extração dos dados, entre eles as *seed categories* mencionadas na seção anterior, porém, em *grounded theory* prega-se que as conclusões devem emergir do conjunto de dados coletados.

- **Sumário Narrativo**

O sumário narrativo é um dos modos mais simples de fornecer uma análise de dados qualitativos. Consiste em sintetizar as ideias principais abordadas no tema de pesquisa, em texto narrativo, generalizando para o leitor os tópicos abordados durante o estudo, desde a metodologia utilizada até resultados obtidos.

Por se tratar de texto em forma narrativa o resumo narrativo oferece como vantagem a possibilidade do pesquisador trabalhar com análise de dados qualitativos em conjunto com dados quantitativos, fator importante comparado a outras técnicas, visto que caberá ao redator interpretar e sintetizar os principais conceitos advindos de cada parcela dos dados. É muito utilizado para sintetizar resultados de revisões sistemáticas e fornecer o estado da arte do estudo.

- **Meta-etnografia**

A meta-etnografia é uma das técnicas clássicas e mais utilizadas para análise de dados qualitativos, principalmente pelas ciências humanas por abordar fortemente os aspectos

culturais envolvidos no estudo.

Foi detalhada em 1988 por George W. Noblit e R. Dwight Hare explicitando os conceitos propostos pela meta-etnografia no livro "Meta-Ethnography: Synthesizing Qualitative Studies" que consiste de sete etapas mostradas na Tabela 2.1.

**Tabela 2.1: Etapas da meta-etnografia**

<b>Etapas da meta-etnografia</b>	<b>Processo envolvido</b>
Iniciando o estudo	Identificar um interesse intelectual que a pesquisa qualitativa possa auxiliar
O que é relevante ao objetivo	Procurando os estudos a serem incluídos
Ler os estudos	Rer ler repetidamente os estudos para identificar os conceitos
Correlacionar os estudos	Determinar o relacionamento entre os estudos.
Traduzir os estudos em outro	Comparação entre os conceitos abordados nos estudos correlacionando-os
Sintetizar a tradução	Segunda tradução (nem sempre possível) onde os conceitos podem ser sintetizados.
Expressar a síntese	Relatar os resultados obtidos por meio de síntese da meta-etnografia em forma apropriada ao contexto.

A técnica relatada, apesar de ser uma das primeiras desenvolvidas, ainda é amplamente utilizada por possibilitar um modo sistemático de analisar os dados e imersão total do pesquisador no objeto de estudo.

#### ● **Estudo de caso**

O estudo de caso consiste em obter e extrair dados (geralmente em questionários) e a partir dos dados qualitativos extraídos transformá-los em dados quantitativos para aplicação de técnicas de abordagem quantitativa. Teixeira (2003) destaca que o estudo de caso é muito utilizado em experimentos de Engenharia de Software nos quais são passados questionários aos participantes e destes serão extraídas as conclusões a respeito do assunto.

Esta técnica fornece uma abordagem prática, porém, alguns cuidados devem ser tomados ao aplicá-la. Sua principal vantagem e ao mesmo tempo desvantagem é a transformação dos dados qualitativos em quantitativos. A análise quantitativa possui métodos testados e eficientes para análise, porém, a transformação dos dados qualitativos em quantitativos deve ser feito de modo preciso para evitar distorções no resultado do estudo.

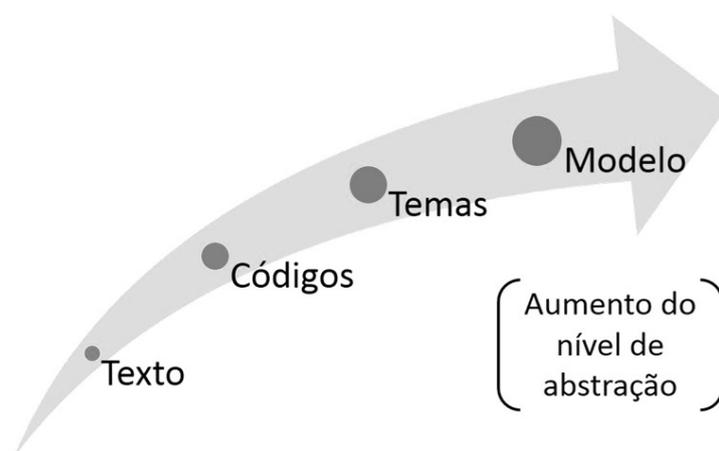
#### ● **Análise temática**

A análise temática consiste em, com base no material de estudo previamente selecionado, analisar, identificar e categorizar os temas emergidos e recorrentes no estudo proposto possibilitando identificar temas chaves para a pesquisa (BOYATZIS, 1998).

Segundo Bardin (1995), é uma das formas que melhor se adequa às investigações qualitativas. Esta técnica consiste em colocar os dados em categorias e determinar a frequência de tais categorias. Como propõe o mesmo autor, três etapas constituem a aplicação desta técnica de análise:

- *Pré-análise*: Nesta etapa, são escolhidos os estudos a serem analisados e é efetuada uma "leitura flutuante" a fim de se imergir no assunto abordado e começar a formulação de hipóteses.
- *Exploração do material*: Os materiais obtidos serão cuidadosamente analisados e os trechos importantes serão destacados e codificados em temas ou categorias.
- *Tratamento dos resultados e interpretação*: A partir dos códigos ou temas gerados serão formuladas as hipóteses e conclusões do assunto abordado a fim de oferecer uma interpretação dos resultados obtidos.

A construção de temas agregando os assuntos possibilita uma visão mais ampla e simplificada do tópico tratado, identifica lacunas na pesquisa e tópicos recorrentes possíveis de trabalhos futuros. De modo simplificado, consiste de, a partir dos dados de entrada, refinar o modelo aumentando o nível de abstração dos dados em informação (Figura 2.15).



**Figura 2.15: Análise temática - adaptada de Cruzes e Dyba (2011)**

Mais uma vez vale ressaltar que, seja qual for as técnicas utilizadas para condução da análise dos dados qualitativos, este processo exige dedicação e trabalho árduo do analista. Uma das principais diferenças entre as abordagens qualitativa e quantitativa é: na abordagem quantitativa

os dados são, geralmente, utilizados para confirmar ou confrontar uma hipótese estabelecida por meio dos dados quantificáveis. Já na abordagem qualitativa, a teoria emerge a partir da análise dos dados.

A utilização das abordagens quantitativa e qualitativa podem ser utilizadas em conjunto. São métodos complementares que contribuem para o desenvolvimento e confirmação de teorias, motivo pelo qual mecanismos para auxiliar a condução de tais são encorajados.

O método qualitativo também vem ganhando espaço nas pesquisas da área de computação, principalmente nas áreas de Engenharia de Software e Interação Humano-Computador (IHC). Este crescimento está relacionado a estas áreas envolverem um misto de aspectos técnicos e humanos. Maxwell (2012) cita que, quando um pesquisador ou avaliador deseja estudar questões que não são facilmente particionadas em entidades discretas, ou para analisar a dinâmica de um processo ao invés de suas características estáticas, métodos qualitativos são mais úteis do que as abordagens quantitativas. Visto que, muitas vezes deve-se analisar a interação do usuário com novas ferramentas ou metodologias, métodos qualitativos podem ser mais indicados nestes casos.

### 2.3.3 Ferramentas para auxiliar a análise qualitativa de dados

Conforme mencionado nas seções anteriores, a análise qualitativa dos dados pode trazer informações valiosas para a pesquisa.

Neste contexto, existem algumas ferramentas que auxiliam a extrair dados dos documentos e analisar os materiais de estudo. A seguir são descritas algumas destas ferramentas disponíveis. As mencionadas ferramentas utilizam como base a aplicação da técnica de *coding*, descrita anteriormente neste trabalho:

- *Nvivo*: É a ferramenta mais famosa e utilizada por pesquisadores. É um software proprietário comercial, produzido pela empresa QSR International.

Permite a análise com base em texto e arquivos multimídia. Além dos recursos de análise de mídia, possui extensões para os navegadores Google Chrome e Internet Explorer para aquisição de material de análise direto dos navegadores.

Assim como as demais ferramentas, está estruturada seguindo a técnica de *coding*.

- *MaxQDA*: Desenvolvida desde 1989 pela VERBI Software - Consult - Sozialforschung GmbH, com sede em Berlim - Alemanha. O processo de análise realizado com o auxílio da ferramenta é similar aos outros: está embasado na técnica de codificação com





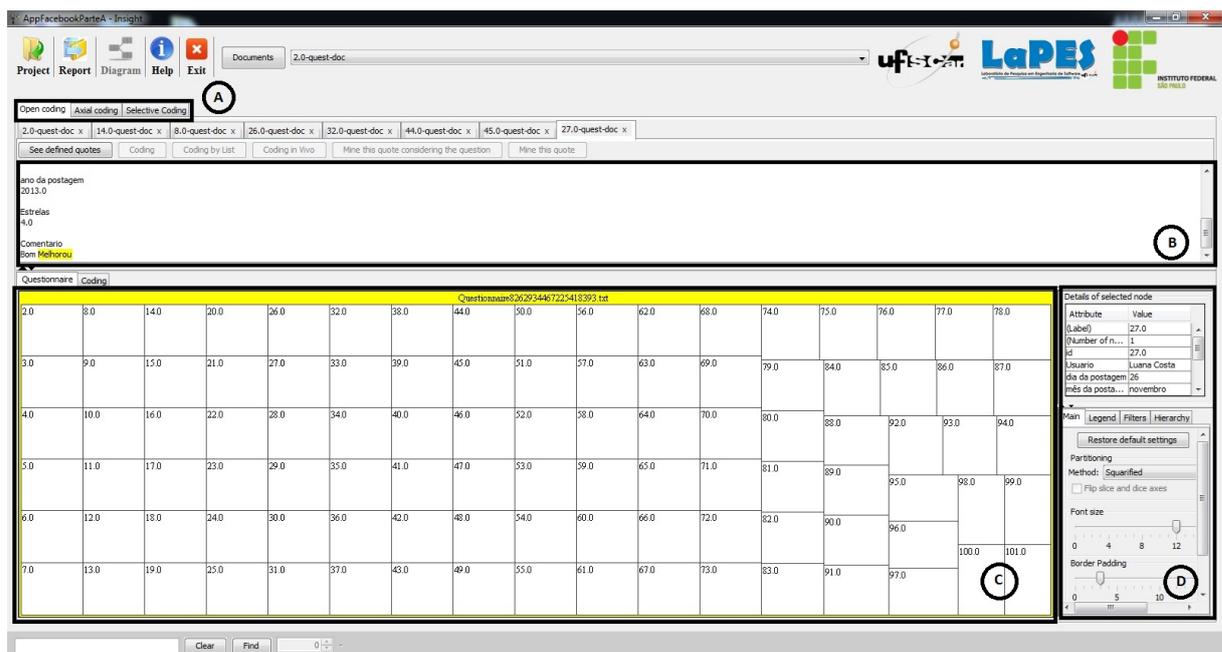
As mencionadas técnicas, quando utilizadas em conjunto, agregam importantes recursos para se conduzir a análise qualitativa de dados pois permitem a abstração de parte dos dados contidos nos documentos. Conforme já descrito no Capítulo 2, a análise qualitativa tem como base o caráter interpretativo da pesquisa por parte do pesquisador. Em vista desta característica intrínseca à análise qualitativa, as técnicas computacionais empregadas na ferramenta contribuem para o pesquisador abstrair informações presentes nos dados.

No que tange o aspecto funcional da ferramenta, ela está estruturada seguindo a técnica de *coding*, conforme descrito por Seaman (2008). O primeiro passo para utilização da ferramenta é a criação de um novo projeto. Para tal, o analista preenche as informações básicas para documentação (autor(es) da pesquisa, breve descrição sobre os objetivos), idioma dos documentos a serem analisados (a mineração de texto implementada é sensível ao idioma) e tipo dos documentos a serem analisados (formato texto ou planilha de dados).

Após criado o projeto é gerado um arquivo com a extensão *.script*, no qual todos os dados relativos à análise serão armazenados. Os documentos a serem analisados devem então ser importados para a ferramenta onde será iniciado o processo de codificação dos trechos. A Figura 2.19 apresenta a interface principal da ferramenta rotulada com os principais componentes, descritos a seguir:

- A** : Aba de navegação entre as etapas de coding (*Open Coding, Axial Coding e Selective Coding*).
- B** : Área de visualização do conteúdo do documento corrente da análise.
- C** : Área de visualização do conjunto de documentos da análise. Cada quadrado representado nesta área corresponde a um documento, no caso de projetos com documentos texto, ou a uma linha da planilha de dados, no caso de projeto em que o documento importado seja uma planilha de dados.
- D** : Nesta área são disponibilizadas as opções para visualização dos dados. Entre as opções disponíveis estão a aplicação de legenda nos documentos, filtro de dados, agrupamento de documentos com base em um atributo, entre outras que serão descritas no decorrer desta seção.

Uma vez importado os documentos da análise na ferramenta, o pesquisador pode utilizar os recursos disponibilizados. A base do processo definido na ferramenta é a codificação dos documentos. Podem ser codificados vários trechos de um mesmo documento, basta selecionar o trecho de interesse e atribuir um rótulo à ele.



**Figura 2.19: Organização da ferramenta Insight**

A ferramenta Insight disponibiliza três opções para atribuição do rótulo ao trecho, conforme a Figura 2.20, relativa a um projeto de análise de *feedback* de usuários do aplicativo móvel Facebook, com os botões destacados com borda preta grossa. Pode-se utilizar a opção "Coding", na qual o usuário cria um novo rótulo e atribui o trecho selecionado a ele. Outra opção, utilizada na Figura 2.20, é a "Coding by List"; o usuário atribui o trecho selecionado a um rótulo criado anteriormente. Também há a possibilidade de utilizar o "Coding in Vivo", ou seja, o rótulo atribuído ao trecho é exatamente o trecho selecionado.

Além do aspecto de codificação implementado na ferramenta, outro recurso disponível e mencionado no início desta seção é a visualização dos dados. Na Figura 2.20, os documentos forma coloridos de acordo com a avaliação dos usuários em relação ao aplicativo móvel da análise (em uma escala de 1 a 5, conforme legenda na ferramenta situada na parte inferior direita).

Com aplicação conjunta à visualização de dados, a ferramenta Insight possui recursos de mineração de texto para identificação da similaridade entre documentos. A mineração de texto implementada na ferramenta utiliza o algoritmo da janela deslizante, descrito por Hernandez et al. (2014). A partir de um trecho selecionado, é calculado um índice de similaridade de cada documento em relação ao trecho e, em conjunto com a visualização de dados, os documentos são coloridos de acordo com legenda disponibilizada na parte inferior da ferramenta (Figura 3.11).

As opções responsáveis pela mineração de texto na Insight estão destacados em borda

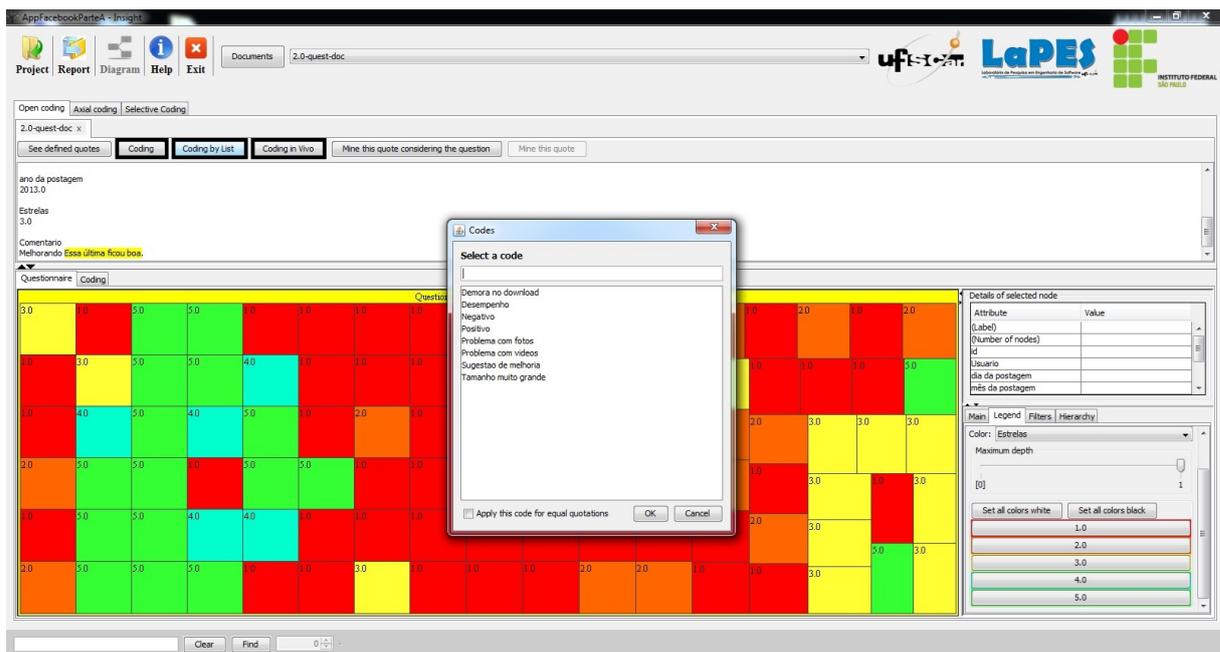


Figura 2.20: Visualização dos dados e codificação dos documentos

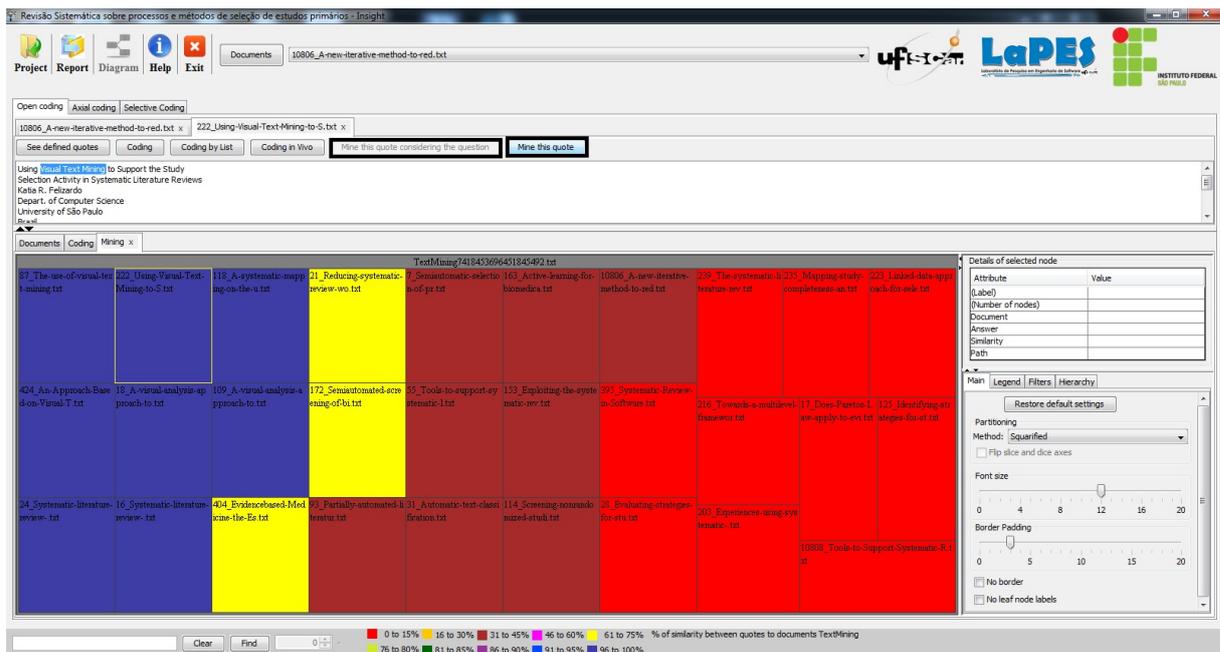


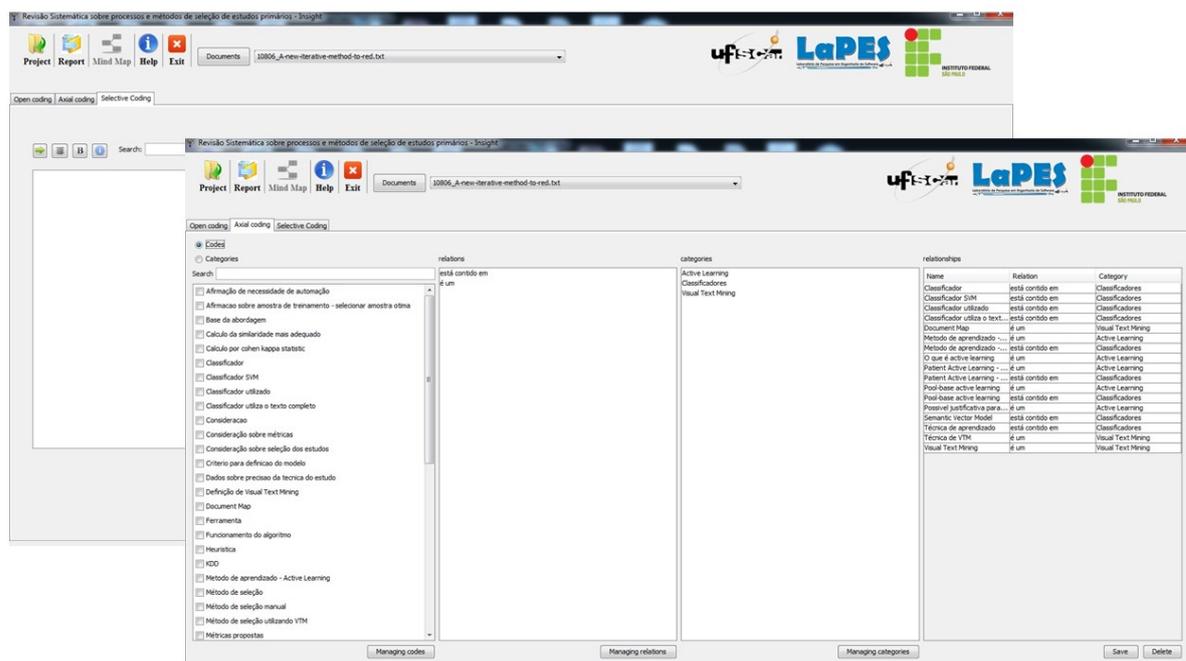
Figura 2.21: Mineração de texto em conjunto com a visualização de dados

grossa preta na Figura 3.11. Assim como a opção de codificação, é possível realizar a mineração de dois modos distintos. Uma delas é utilizar a opção "Mine this quote considering the question", aplicável apenas para análises de planilha de dados. Com esta opção, a mineração é aplicada com base no trecho selecionado e apenas na coluna equivalente da planilha para cada entrada de dados. A outra opção disponível, "Mine this quote", é aplicável apenas para projeto com entrada de documentos texto - efetua a mineração baseada em todo o conteúdo dos

documentos.

As opções descritas acima correspondem a recursos para auxiliar, principalmente, a etapa de *Open Coding* do processo da técnica *Coding*. Como mencionado no início desta seção, a ferramenta Insight está estruturada seguindo a técnica *Coding* e dá suporte à aplicação de todas as etapas da mesma. Na Figura 2.22, é apresentado no primeiro plano o suporte da ferramenta para o *Axial Coding* e, em segundo plano, para o *Selective Coding*.

Para a etapa de *Axial Coding*, são listados todos os códigos que o usuário atribuiu durante a fase de *Open Coding*. A partir destes, é possível criar categorias, agrupar os códigos nas categorias e definir os relacionamentos que os códigos e categorias possuem uns com os outros. Todos os relacionamentos, códigos e categorias são listados na própria ferramenta para o usuário poder exportá-los em relatórios ou mapas mentais. Para realização do *Selective Coding*, a ferramenta Insight disponibiliza um editor de texto para sumarização dos resultados.



**Figura 2.22: Etapas de Axial e Selective Coding na ferramenta Insight**

Por fim, é possível exportar os dados coletados da análise para mapas mentais ou relatórios de extensão .xls. A janela destacada com borda preta na Figura 2.23 ilustra as opções de possíveis relatórios a serem gerados. Podem ser selecionados conjuntos de dados a serem exportados ou de todos os dados extraídos para um conjunto de documentos ou então para todos os documentos.

Nos parágrafos anteriores desta seção foram descritos as principais funcionalidades da ferramenta Insight. Quanto aos aspectos de implementação da ferramenta, assim como a StArt,

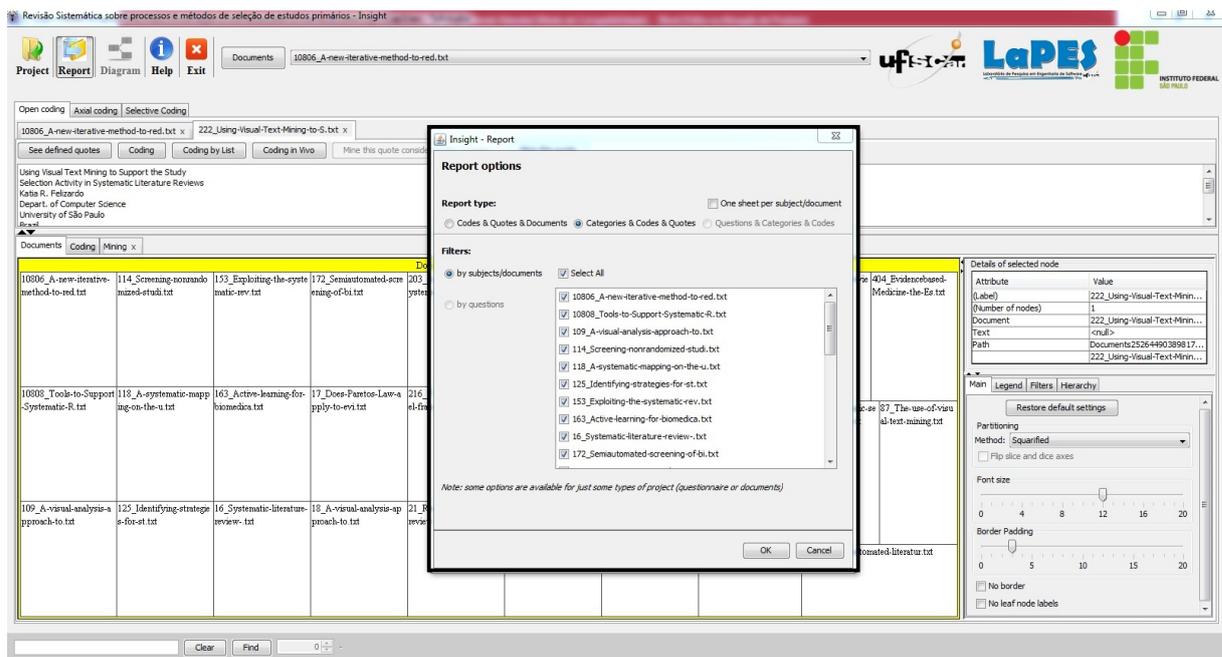


Figura 2.23: Exportação dos relatórios da análise

foi desenvolvida utilizando linguagem de programação Java. O armazenamento dos dados é executado com o gerenciador de banco de dados HSQLDB em modo arquivo, devido a sua portabilidade, com a persistência no banco utilizando a API (*Application Programming Interface*) JPA. Entre demais componentes utilizados estão: biblioteca APACHE POI para interpretação de planilhas de dados como entrada de dados da ferramenta e a biblioteca Jasper Reports para geração dos relatórios relativa à análise realizada.

Vale ressaltar também entre os aspectos de implementação, a utilização da ferramenta Treemap, fornecido pela universidade de Maryland ao LAPES. A integração entre as ferramentas foi realizada com a implementação de um módulo que integrou-a à Insight para auxiliar na visualização dos dados.

## 2.4 Mapeamento sistemático sobre utilização de análise qualitativa em revisões sistemáticas

Como ponto de partida para esta pesquisa, foram realizados estudos sobre técnicas de análise qualitativa de dados - entre elas meta-etnografia, análise temática e outras. Também investigou-se os tipos de estudos secundários e como são conduzidos.

Após compreensão das técnicas de análise qualitativa, descritas na seção 2.3.2, foi conduzido um mapeamento sistemático a fim de identificar de que modo é utilizada a análise qualita-

tiva em revisões sistemáticas

### Planejamento e condução

A Tabela 2.2 sumariza o planejamento do mapeamento conduzido.

**Tabela 2.2: Planejamento do mapeamento sistemático conduzido**

<b>Analisar</b>	Utilização de análise qualitativa em revisão sistemática
<b>Com o propósito de</b>	Mapear a área de estudo.
<b>Com respeito a</b>	Ferramentas e processos utilizados
<b>Do ponto de vista do</b>	Pesquisador
<b>No contexto de</b>	Estudos que utilizam as técnicas em conjunto

A seguir, são detalhados os principais aspectos relativos ao planejamento e condução do mapeamento conduzido neste trabalho.

- **Descrição do Problema:** Em uma revisão sistemática o volume de dados obtidos é muito grande pois se trata de uma busca extensa na literatura. A aplicação da análise qualitativa vai ao encontro da necessidade de técnicas de extração e sumarização dos dados obtidos da revisão sistemática a fim de disponibilizar o estado da arte do tema abordado.
- **Objetivo:** O foco deste mapeamento foi levantar um panorama geral do contexto em que está sendo utilizada análise qualitativa em conjunto com revisão sistemática. Foram identificadas técnicas e ferramentas que vêm sendo utilizadas na área citada.
- **Questões de Pesquisa:** As questões de pesquisa respondidas no estudo foram:
  - **Questão Principal:** "Qual o panorama do uso de análise qualitativa em conjunto com a revisão sistemática?"
  - **Questões secundárias:**
    - Q1: Qual o local de origem do estudo?
    - Q2: Qual a área de estudo?
    - Q3: Quais técnicas foram utilizadas para a análise qualitativa?
    - Q4: Qual ferramenta foi utilizada para auxiliar o processo?
- **Palavras-chave:** As palavras chave utilizadas para compor a string de busca nas bases de dados foram as seguintes: análise qualitativa, *qualitative analysis*, revisão sistemática, *systematic review*, *systematic literature review*.

- **Método utilizado para pesquisa de fontes primárias:** Como o mapeamento busca uma visão abrangente sobre as publicações mais recentes do tópico pesquisado, o método utilizado para pesquisa de fontes primárias foi a busca em bases de dados eletrônicas.
- **Critérios para inclusão e exclusão dos estudos:** Para realizar o mapeamento foram estabelecidos critérios para exclusão ou inclusão dos estudos. Somente foram incluídos os que atenderam ao critério de inclusão e não atenderam a nenhum critério de exclusão. Os critérios utilizados são listados a seguir:
  - *Inclusão:* Efetuou uma análise qualitativa em conjunto com revisão sistemática.
  - *Exclusão:* Não efetuou análise qualitativa.
  - *Exclusão:* Não efetuou revisão sistemática.
  - *Exclusão:* Sem acesso ao artigo completo.
- **Crítério de qualidade dos estudos:** Na primeira etapa da pesquisa, a qual consistiu em obter os dados requisitados no formulário de extração definido, a qualidade dos estudos não foi avaliada. Na segunda etapa, mais aprofundada e investigando como o processo de revisão sistemática em conjunto com análise qualitativa era conduzido, os estudos foram avaliados para verificar se descreviam detalhadamente como o processo foi realizado.
- **Método de avaliação dos estudos:** Os estudos foram avaliados individualmente seguindo os critérios de inclusão e exclusão. A avaliação se concentrou nos tópicos sobre quais metodologias foram utilizadas para os estudos em questão devido ao foco do mapeamento ser identificar a aplicação de revisão sistemática em conjunto com análise qualitativa.
- **Método de extração dos dados:** Os dados foram extraídos do seguinte modo:
  - Os potenciais estudos primários foram coletados e importados para a ferramenta *Start*.
  - Foram aplicados os critérios de seleção a todos os estudos efetuando a leitura dos *abstracts*.
  - Os estudos aceitos na etapa de seleção foram submetidos a etapa de extração de dados. Os dados foram extraídos utilizando os formulários de extração previamente definidos no protocolo do mapeamento.
  - A partir dos formulários de extração os dados foram exportados pela ferramenta *Start* para o Microsoft Office Excel 2013 para condução da análise e extração dos dados.

- **Método de síntese dos dados:** Foi realizada uma meta-análise sobre os dados quantitativos extraídos dos estudos analisados.

De modo simplificado, o mapeamento sistemático conduzido teve por objetivo identificar o panorama geral do contexto em que a análise qualitativa é utilizada em conjunto com a revisão sistemática e qual a técnica de análise qualitativa mais adequada quando aplicada em conjunto com a revisão sistemática.

### Dados coletados

Nos tópicos a seguir são apresentados os dados coletados do mapeamento para cada questão de pesquisa definida:

- **Q1: Qual o local de origem do estudo?**

O objetivo desta questão 1 foi identificar os países onde as publicações a respeito do tema deste mapeamento estão sendo escritas, os resultados podem ser vistos na Tabela 2.3.

**Tabela 2.3: Países de publicação**

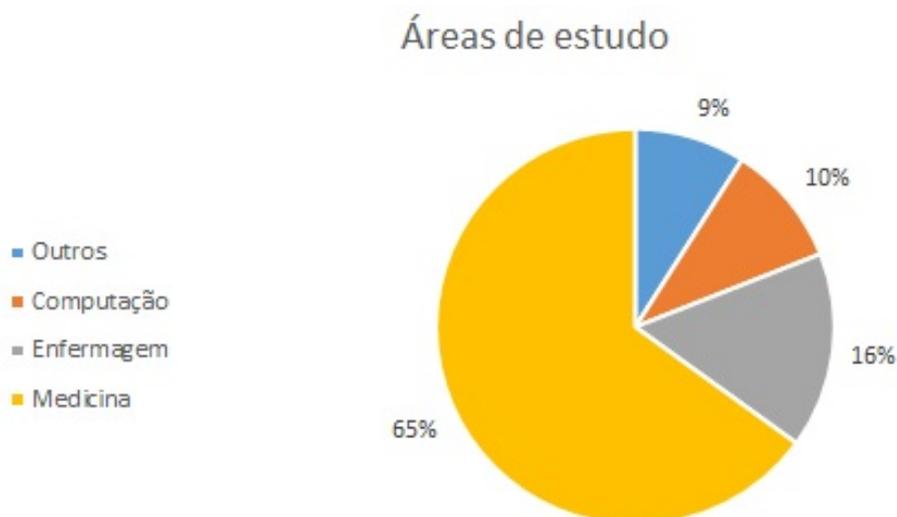
País de publicação	Quantidade	País de publicação	Quantidade
Inglaterra	64	Alemanha	3
Austrália	19	Suíça	2
EUA	12	País de Gales	2
Canadá	9	Itália	2
Noruega	7	Irlanda	2
Holanda	5	Espanha	2
Nova Zelândia	4	Uruguai	1
China	3	Suécia	1
Brasil	3	Portugal	1

- **Q2: Qual a área de estudo?**

O objetivo da questão de pesquisa 2 foi identificar as áreas onde as técnicas estão sendo utilizadas em conjunto (Figura 2.24).

Os estudos da área de medicina representam a maioria visto que o processo de revisão sistemática é prática consolidada na área e dados qualitativos são frequentes devido ao trabalho realizado envolver por diversas vezes fatores humanos e sociais como material de estudo.

Estudos de computação, mais precisamente de engenharia de software, vem recentemente incorporando dados qualitativos às análises realizadas e aparecem com quantidade considerável de trabalhos conduzidos.



**Figura 2.24: Áreas de estudo**

• **Q3: Quais técnicas foram utilizadas para a análise qualitativa?**

Entre os estudos analisados foram identificadas as seguintes técnicas para conduzir a análise qualitativa dos dados:

- *Análise temática*: Utilizada na maioria dos estudos (exceto um que utilizou meta-etnografia).
- *Meta-Etnografia*: Foi utilizada em um dos estudos. É uma das técnicas clássicas de análise qualitativa de dados.

• **Q4: Qual ferramenta foi utilizada para auxiliar o processo?**

Neste mapeamento buscamos também identificar quais ferramentas foram utilizadas no contexto de análise qualitativa em estudos secundários - e se foi utilizada alguma ferramenta. Dos 138 estudo primários selecionados, 31 mencionaram a utilização de alguma ferramenta (Tabela 2.4).

Entre as ferramentas utilizadas, a mais citada foi a Nvivo - ferramenta especialista para análise qualitativa de dados. Além da Nvivo, outras ferramentas para análise qualitativa de dados foram citadas, entre elas HyperResearch, MaxQDA e Atlas.ti. Além de ferramentas para análise qualitativa, foram citados gerenciadores de referência (EndNote e Reference Manager) e outras com propósitos não específicos para as técnicas (Excel).

**Análise de trabalhos utilizando as técnicas de revisão sistemática e análise qualitativa em conjunto**

Além dos dados brutos coletados, foi selecionado um conjunto de estudos primários que

**Tabela 2.4: Ferramentas utilizadas**

<b>Ferramenta</b>	<b>Quantidade</b>	<b>Ferramenta</b>	<b>Quantidade</b>
<b>Nvivo</b>	8	<b>CHIP Tool</b>	1
<b>Excel</b>	7	<b>QDA Miner Software</b>	1
<b>HyperResearch</b>	4	<b>Reference Manager</b>	1
<b>Casp Tool</b>	3	<b>RevMan</b>	1
<b>MaxQDA</b>	3	<b>StatsDirect</b>	1
<b>EndNote</b>	2	<b>Word</b>	1
<b>Atlas.ti</b>	1		

descreviam o modo como utilizaram as técnicas em conjunto. Verificamos que o uso das técnicas em conjunto é uma prática relativamente nova. Encontrar artigos que detalham como foi realizado o processo por completo não é uma tarefa simples.

Nos tópicos que seguem serão apresentados artigos que utilizaram as técnicas de revisão sistemática e análise qualitativa de dados em conjunto. Para cada artigo foram destacados o título do estudo, autores que conduziram o processo e resumo do tópico abordado no artigo bem como as informações principais de como o processo de revisão sistemática foi utilizado em conjunto com a análise qualitativa de dados.

- *Adolescent views on transition in diabetes and nephrology* (TONG et al., 2013)

**Autores:**Allison Tong, Germaine Wong, Elisabeth Hodson, Rowan G. Walker, Lidwien Tjaden, Jonathan C. Craig.

**Análise do artigo:** O objetivo deste artigo foi sintetizar estudos a respeito da transição de jovens com diabetes ou doença crônica no rim para a fase adulta. A pesquisa foi conduzida por dois pesquisadores em conjunto. Foi realizada uma revisão sistemática de pesquisas e estudos qualitativos que exploravam o tema.

Para os questionários quantitativos, os itens foram mapeados em uma tabela. Já para os estudos qualitativos, foi efetuada a análise temática na qual foram extraídas as anotações dos participantes e os textos das seções de conclusão/discussão - esses dados extraídos foram inseridos textualmente no software HyperResearch.

Para cada artigo, foram codificados os trechos relativos ao tema em processo repetido mais de uma vez. Ao elaborar a conclusão foi usada a técnica de mapa mental para correlacionar os temas obtidos (ao final foram identificados cinco temas principais).

- *Behavioral interventions for weight management in pregnancy: A systematic review of quantitative and qualitative data* (CAMPBELL et al., 2011)

**Autores:** Fiona Campbell, Maxine Johnson, Josie Messina, Louise Guillaume, Elisabeth Goyder.

**Análise do artigo:** Foi realizada uma revisão sistemática de evidências quantitativas e qualitativas. Incluiu uma meta-análise utilizando a ferramenta RevMan (*Review Manager*) e uma análise temática de estudos qualitativos que investigaram as opiniões das mulheres sobre a gestão de peso durante a gravidez.

Para extração dos dados, foram utilizados formulários diferentes para os estudos qualitativos e quantitativos. Os formulários foram desenvolvidos por especialistas da área da pesquisa.

A síntese dos dados foi realizada em três passos conforme descrito por Thomas et al. (2004):

1. Em primeiro lugar, sempre que possível e, se for o caso, os resultados de estudos controlados elegíveis foram estatisticamente sintetizados em uma meta-análise para avaliar a eficácia das intervenções nos ensaios controlados. Após, os estudos foram divididos em categorias pré-especificadas.
2. Em segundo lugar, uma análise temática dos resultados dos estudos qualitativos foi realizada. Cada estudo foi lido e re-lido para permitir que ao revisor familiarizar-se com os resultados do estudo e os métodos utilizados. Os resultados do estudo foram codificados linha por linha para caracterizar o conteúdo de cada linha ou frase. Após resumir os resultados foram identificados temas comuns e apoiar citações extraídas dos estudos qualitativos.
3. Em terceiro lugar, uma matriz metodológica e conceitual foi construída para integrar os resultados das duas sínteses.

Após os dados qualitativos e quantitativos alinhados, as conclusões finais a respeito do tema foram obtidas pelos autores analisando os dados em conjunto.

- *Empirical studies of agile software development: A systematic review* (DYBÅ; DINGSØYR, 2008)

**Autores:** Tore Dybå, Torgeir Dingsøy.

**Análise do artigo:** Neste artigo foi realizada uma revisão sistemática sobre desenvolvimento ágil de software. Foram analisadas pesquisas qualitativas e quantitativas. Os estudos foram avaliados por dois pesquisadores seguindo 11 critérios de qualidade (os focos principais foram avaliação da qualidade da pesquisa qualitativa e por princípios de boas práticas para realização de pesquisa empírica em engenharia de software).

Para extração dos dados foi utilizado um formulário pré-definido. Como os dados foram extraídos por dois pesquisadores separadamente, ocorreram discrepâncias nos resultados extraídos, tais discrepâncias foram resolvidas através de reunião de consenso.

Os objetivos, definições, descrições, métodos de pesquisa, descobertas e conclusões, conforme relatado pelos autores dos estudos primários, foram copiados integralmente para o software NVivo (software especialista para análise qualitativa dos dados textuais).

Foi utilizada a técnicas de análise temática para agrupar os estudos nos seguintes temas: (1) Introdução e adoção, (2) Humanas e fatores sociais, (3) Percepções sobre métodos ágeis e (4) Estudos comparativos.

- *Features of effective primary health care models in rural and remote Australia: a case-study analysis* (WAKERMAN et al., 2009)

**Autores:** John Wakerman, John S Humphreys, Robert Wells, Pim Kuipers, Judith A Jones, Philip Entwistle, Leigh Kinsman.

**Análise do artigo:** Este artigo utilizou uma revisão sistemática realizada anteriormente como início de sua pesquisa para identificar o foco do estudo detalhado a ser realizado.

Foram selecionados seis temas principais. A partir deles, a pesquisa foi aprofundada realizando entrevistas com informantes chave para o processo.

As entrevistas realizadas foram semi-estruturadas, gravadas e posteriormente transcritas. Dois entrevistadores documentaram os temas e resumiram as entrevistas com base nas gravações obtidas.

Embora as entrevistas cobrissem diversas questões, temas claramente consistentes surgiram na análise inicial da pós-entrevista e posterior codificação e análise independente.

A partir dos textos que foram obtidos das entrevistas, foi realizado o processo de codificação dos trechos e geração dos temas. Para validar o processo e identificar os temas excluídos, as transcrições foram tematicamente codificadas por outro membro da equipe que não esteve envolvido nas entrevistas utilizando o software Nvivo.

- *How patients understand depression associated with chronic physical disease? A systematic review* (ALDERSON et al., 2012)

**Autores:** Sarah L Alderson, Robbie Foy, Liz Glidewell, Kate McLintock, Allan House.

**Análise do artigo:** Este artigo foi desenvolvido a fim de compreender os sintomas e causas da depressão em pacientes. Foi conduzida uma revisão sistemática da literatura

em conjunto com o processo de análise temática. O processo foi conduzido por dois pesquisadores em conjunto.

A partir dos estudos incluídos na revisão, formulários de extração desenvolvidos para esta pesquisa especificamente foram utilizados para coletar os dados considerados relevantes ao contexto. Os resultados obtidos foram inseridos no software de análise qualitativa de dados Nvivo para codificação e, a partir dos códigos gerados, gerar e correlacionar os temas obtidos.

Para sumarizar os resultados da revisão os autores utilizaram a técnica de sumário narrativo, justificando esta escolha pela pesquisa apresentar grande variação no tipo de estudo incluído.

- *I'll be working my way back: A qualitative synthesis on the trauma experience of Children* (WESEL et al., 2012)

**Autores:** Floryt van Wesel, Hennie Boeije, Eva Alisic, Sarah Drost.

**Análise do artigo:** O artigo efetua a busca por estudos primários em bases de dados e, através dos passos definidos da revisão sistemática, chega a um total de 17 pesquisas qualitativas para serem analisadas a fim de obter teorias a respeito do desenvolvimento de estresse pós-traumático, especialmente em crianças.

A síntese qualitativa desenvolvida consistiu de cinco etapas, desenvolvida por dois pesquisadores: (1) busca de estudos primários, (2) avaliação da qualidade dos estudos, (3) extração dos resultados, (4) análise dos resultados e (5) síntese dos resultados.

Os estudos obtidos foram codificados em relação aos métodos e amostras utilizadas.

Cada artigo foi lido cuidadosamente e, para cada, foi escrito um resumo sobre o mesmo. Após os resumos serem elaborados, os mesmos foram comparados e foram notadas semelhanças.

Os resultados obtidos considerados relevantes foram sintetizados utilizando o que é referido por vezes como análise de meta-dados (PATERSON; CANAM, 2001). O método utilizado para a análise é similar a análise temática.

O software de análise qualitativa MaxQDA foi utilizado para auxiliar o processo. Os resumos codificados foram comparados e, a partir dos códigos gerados, foram feitos os correlacionamentos entre eles.

Após efetuada as comparações, os códigos foram agrupados por dois modos: (1) Baseado na frequência em que os temas apareceram nos estudos e (2) Baseado em áreas a que os temas diziam a respeito (chamado pelos autores de domínio).

- *Patients' views on the impact of stroke on their roles and self: A thematic synthesis of qualitative studies* (SATINK et al., 2013)

**Autores:** Ton Satink, Edith H. Cup, Irene Ilott, Judith Prins, Bert J. de Swart, Maria W. Nijhuis-van der Sanden.

**Análise do artigo:** O artigo efetua uma revisão sistemática na qual estudos qualitativos sobre perspectivas de pessoas pós acidente vascular cerebral (AVC). A questão principal que o estudo busca responder é "Como fazer as pessoas com AVC visualizar o impacto de acidente vascular cerebral em seus papéis e eu?".

Os autores relatam que sínteses de pesquisas qualitativas são uma forma bem estabelecida de revisão sistemática da literatura que reúne resultados de estudos qualitativos para contribuir com novas compreensões de perspectivas dos pacientes.

O método utilizado consistiu em quatro etapas principais: (a) Pesquisar na literatura estudos primários relevantes; (b) Avaliação da qualidade dos estudos; (c) Extração dos dados e (d) Geração de temas descritivos e analíticos sobre o assunto abordado.

A etapa de extração aplicou a técnica de coding. Os trechos relevantes dos textos foram rotulados como "resultados" ou "descobertas" e, após, os dados foram importados para o software de análise qualitativa de dados Atlas.ti.

A análise dos dados e a síntese consistiu em três etapas. Primeiramente os dados extraídos foram lidos como um todo e os segmentos significativos foram novamente codificados. A próxima etapa consistiu em organizar os segmentos codificados em áreas para gerar os temas descritivos. Finalmente, os temas descritivos foram discutidos e os temas mais abstratos e analíticos foram desenvolvidos.

O processo todo envolveu o desenvolvimento de interpretações além dos estudos primários. Um mapeamento conceitual apoiou a discussão e o processo de geração dos temas descritivos e analíticos.

- *Quality of life and leisure participation in children with neurodevelopmental disabilities: a thematic analysis of the literature* (DAHAN-OLIEL; SHIKAKO-THOMAS; MAJNEMER, 2012)

**Autores:** Noemi Dahan-Oliel, Keiko Shikako-Thomas, Annette Majnemer.

**Análise do artigo:** Este artigo efetua uma revisão sistemática sobre crianças com deficiência do desenvolvimento neurológico e utilizou a análise temática para obter os resultados e gerar temas de estudo.

Para efetuar a revisão foram incluídos estudos qualitativos, quantitativos e métodos mistos. Para cada tipo de estudo um tipo específico de análise de qualidade foi utilizado. Para os estudos quantitativos e e de métodos mistos, apenas os dados de medidas diretamente relacionados a qualidade de vida foram considerados, fato que impediu uma meta-análise formal como método de avaliação.

A partir dos dados obtidos, foi conduzida uma análise temática para sintetizar as relações obtidas da pesquisa. Os autores justificam a escolha pela técnica por ser uma metodologia qualitativa apropriada para revisões de literatura, pois proporciona uma síntese dos dados originais qualitativos, quantitativos e/ou de métodos mistos através da extração de temas e sub-temas.

As seções resultados de todos os estudos selecionados foram independentemente codificados manualmente para gerar as categorias iniciais que organizavam os dados em temas significativos para o contexto.

- *The Motivations and Experiences of Living Kidney Donors: A Thematic Synthesis* (TONG et al., 2012)

**Autores:** Allison Tong, Jeremy R. Chapman, Germaine Wong, John Kanellis, Grace McCarthy, Jonathan C. Craig.

**Análise do artigo:** O artigo teve como objetivo sintetizar estudos qualitativos sobre as experiências e perspectivas de pessoas doadoras de rins vivas. A revisão sistemática foi conduzida por dois pesquisadores a fim de obter os estudos a serem avaliados. Para inclusão foram aceitos apenas estudos qualitativos.

Os dados foram extraídos através de formulários desenvolvidos especificamente para este propósito, onde eram obtidas informações como critérios utilizados para a pesquisa, métodos e contexto do estudo, análise e interpretação.

A técnica de análise temática foi utilizada nos dados extraídos. Para cada artigo as citações de participantes e textos sobre Resultados/Conclusões foram extraídos e inseridos na íntegra no software HyperResearch.

Os materiais foram codificados linha a linha para geração dos temas descritivos e analíticos para desenvolvimento. Os temas gerados foram correlacionados e, ao final do estudo, um total de seis temas a serem abordados foram desenvolvidos.

### **Sumarização**

Em suma, os trabalhos efetuam a revisão sistemática manualmente de acordo com passos pré-determinados: busca do material de estudo, seleção dos artigos, extração e sumarização dos

dados. Durante a extração dos dados é iniciado o processo de análise qualitativa - geralmente utilizando a técnica de *coding*.

Outro fator relevante observado neste mapeamento, foi a falta de suporte computacional utilizado para condução dos processos. Como mencionado, apenas 31 estudos citaram a utilização de ferramentas. Das ferramentas utilizadas destacam-se gerenciadores de referências e ferramentas específicas para análise qualitativa.

A partir da análise dos estudos e da pesquisa por técnicas de análise qualitativa conduzida anteriormente, verificamos que a técnica de análise qualitativa mais adequada para ser aplicada em revisão sistemática é a análise temática. A análise temática como método de análise de dados auxilia na categorização dos dados, correlacionamento de informações e recuperação destes dados para sumarização. Deste modo, identificamos que auxilia a obtenção do estado da arte em revisões sistemáticas.

A apresentação da abordagem proposta neste trabalho, junto ao suporte computacional desenvolvido, pode auxiliar os usuários que optem pela condução de estudos secundários com análise temática.

## **2.5 Considerações finais**

Neste capítulo foram descritos os conceitos utilizados para elaboração do trabalho para entendimento do leitor bem como sua aplicação na área de Engenharia de Software.

Este trabalho engloba fundamentalmente a utilização da análise qualitativa em conjunto com o processo de revisão sistemática, apoiado pela utilização das ferramentas StArt e Insight. Estas duas técnicas em conjunto podem contribuir para melhoria da qualidade da pesquisa e facilitar o trabalho de extração dos dados e obtenção do estado da arte sobre o tema.

# Capítulo 3

## ANÁLISE TEMÁTICA EM REVISÃO SISTEMÁTICA

---

---

### 3.1 Considerações iniciais

Como mencionado no Capítulo 2, conduzir um estudo secundário ou uma análise de dados com enfoque qualitativo não é uma tarefa simples, demanda tempo, dedicação e planejamento por parte do pesquisador. Nesse sentido a condução dessas tarefas por meio de ferramentas computacionais pode possibilitar uma melhor execução dos estudos realizados.

Assim para a realização deste trabalho foram utilizadas as ferramentas StArt e Insight, as quais foram desenvolvidas pelo Laboratório de Pesquisa em Engenharia de Software da UFS-Car. Visto que a StArt foi desenvolvida com o propósito específico para condução de estudos secundários e a Insight para analisar dados qualitativos, novas funcionalidades foram implementadas nas duas ferramentas a fim de permitir a integração delas por meio do compartilhamento de arquivos.

Este capítulo está organizado conforme descrito a seguir. A seção 3.2 apresenta a caracterização da análise temática em conjunto com estudos secundários. Na seção 3.2.1 é descrita a caracterização da análise temática em conjunto com a revisão sistemática. Na seção 3.2.2 é apresentada a abordagem RS/AT, desenvolvida neste trabalho para aplicação da análise temática em conjunto com revisão sistemática. Finalmente, a seção 3.3, apresenta uma síntese dos desenvolvimentos realizados nas ferramentas StArt e Insight.

## **3.2 Caracterização e desenvolvimento da análise temática em conjunto com estudos secundários**

Quando se realiza um estudo secundário, diversos tipos de estudos primários são coletados. Cada estudo primário pode realizar sua síntese e apresentação das informações de modo distinto. Porém, independentemente do modo que os dados e as informações são apresentadas nos estudos primários, há necessidade de utilização desses estudos na composição e síntese de um estudo secundário. Assim, observa-se a necessidade de técnicas para aumentar a produtividade e qualidade dessas análises.

Tradicionalmente, a extração dos dados para estudos secundários sistemáticos é realizada por meio de formulários de extração. Esses formulários são desenvolvidos pelos pesquisadores envolvidos no estudo e/ou por especialistas na área. Porém, por diversas vezes há a necessidade de, durante a revisão, adicionar ou remover campos definidos nesse formulário. A fim de manter a extração dos dados consistentes entre os estudos, deve-se então analisar os estudos que já haviam sido analisados, a fim de extrair informações para contemplar os novos campos de extração de dados inseridos. Este processo iterativo da revisão sistemática é chamado de I-SR (revisão sistemática incremental) e descrito por Fabbri et al. (2013).

Neste contexto, pretende-se aplicar a abordagem qualitativa, por meio da análise temática, como método auxiliar para a condução de estudos secundários, uma vez que a extração de dados pode ser beneficiada por ela.

### **3.2.1 Caracterização da análise temática em conjunto com a revisão sistemática**

Conforme descrito no Capítulo 2, o mapeamento sistemático possui enfoque quantitativo, buscando identificar o panorama da área de estudo em questão. Já a revisão sistemática busca extrair e sumarizar dados detalhados dos estudos primários a fim de caracterizar o estado da arte do tópico em questão com riqueza de detalhes.

Uma das diferenças entre ambos é o modo como a qualidade do estudo é aferida. O mapeamento é menos criterioso quanto à qualidade do estudo em questão e a extração dos dados requisitados busca chegar ao panorama geral da área pesquisada.

Nesse contexto identificou-se que, em estudos secundários, a aplicação da análise temática poderia apresentar contribuições ao processo quando aplicada em revisões sistemáticas, uma vez que esta implica em um entendimento mais profundo do tema pesquisado. Ressalta-se que

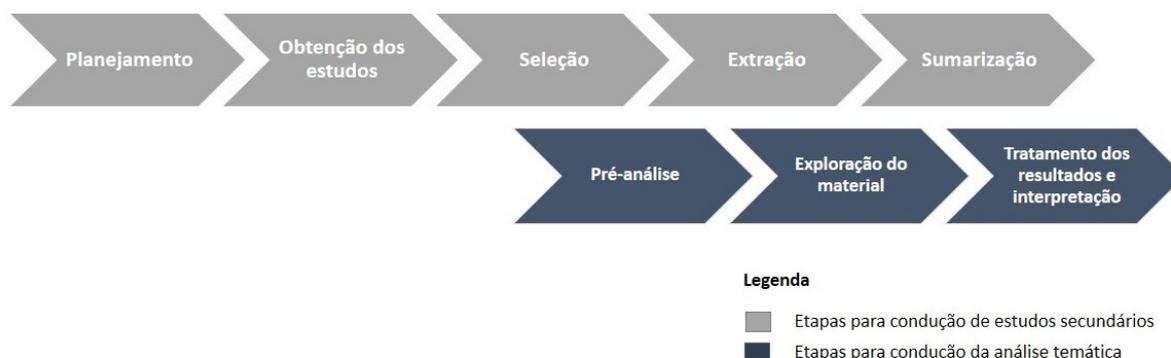
a análise temática deve contribuir para o estabelecimento das categorias a serem definidas em um mapeamento sistemático, mas essa investigação não está no escopo deste trabalho e deve ser feita futuramente.

Com base no mapeamento sistemático descrito no capítulo anterior, que teve por objetivo identificar as técnicas de análise qualitativa utilizadas no processo de revisão sistemática, constatou-se que a análise temática é a técnica mais empregada.

Os pontos propícios à integração dos processos de revisão sistemática e análise temática foram identificados ao se efetuar uma análise detalhada dos artigos considerados no mapeamento que utilizavam essas técnicas (análise temática e revisão sistemática) em conjunto. Essas conclusões foram obtidas visto que, ao analisar as técnicas e pesquisas de análise qualitativa, verificou-se que o enfoque qualitativo explora o tópico em questão com detalhes e busca efetuar correlacionamento entre dados e informações obtidas.

A revisão sistemática é um processo estruturado e com passos bem definidos. Já a análise temática, assim como outras técnicas de análise qualitativa, possui caráter exploratório e investigativo. Uma comparação das etapas realizadas em cada uma das técnicas e como transcorre o fluxo de dados nos processos é apresentado na Figura 3.1.

Na análise temática, a primeira etapa a ser conduzida é a pré-análise. Consiste em efetuar uma leitura flutuante nos estudos para análise a fim de imergir o pesquisador no tema da análise. Para a pré-análise, um dos requisitos é possuir o material de estudo já selecionado. Em revisões sistemáticas, o pesquisador efetua todo o planejamento, obtenção e seleção dos estudos para então iniciar a extração e sumarização dos dados. Nesse contexto, a pré-análise da análise temática pode ser incluída como uma nova etapa em uma abordagem de revisão sistemática em conjunto com análise temática. As demais etapas podem ser correlacionadas umas com as outras (este tópico é detalhado na próxima subseção - 4.2.2).



**Figura 3.1: Fluxos das etapas de revisão sistemática e da análise temática**

O material para estudo nessa área ainda é escasso. Muitos dos trabalhos que utilizam as técnicas de revisão sistemática e análise temática em conjunto, não documentam como o processo foi conduzido.

Basicamente, ao se conduzir a análise temática de um tópico de interesse, os pesquisadores efetuam a busca de estudos primários, avaliam a qualidade dos estudos obtidos, extraem resultados (utilizando o método de preferência) e, por fim, analisam e sintetizam os resultados - como os estudos realizados por Dybå e Dingsøy (2008), Campbell et al. (2011) e Satink et al. (2013).

Apesar de a análise temática não apresentar um processo rigidamente estruturado, ela se relaciona diretamente com as etapas de extração e sumarização dos dados em estudos secundários, conforme mencionado anteriormente. Em vista disso, a próxima seção descreve detalhadamente uma abordagem formulada neste trabalho para aplicação das técnicas em conjunto com auxílio das ferramentas StArt e Insight.

### **3.2.2 Aplicando análise temática em revisão sistemática com a abordagem RS/AT**

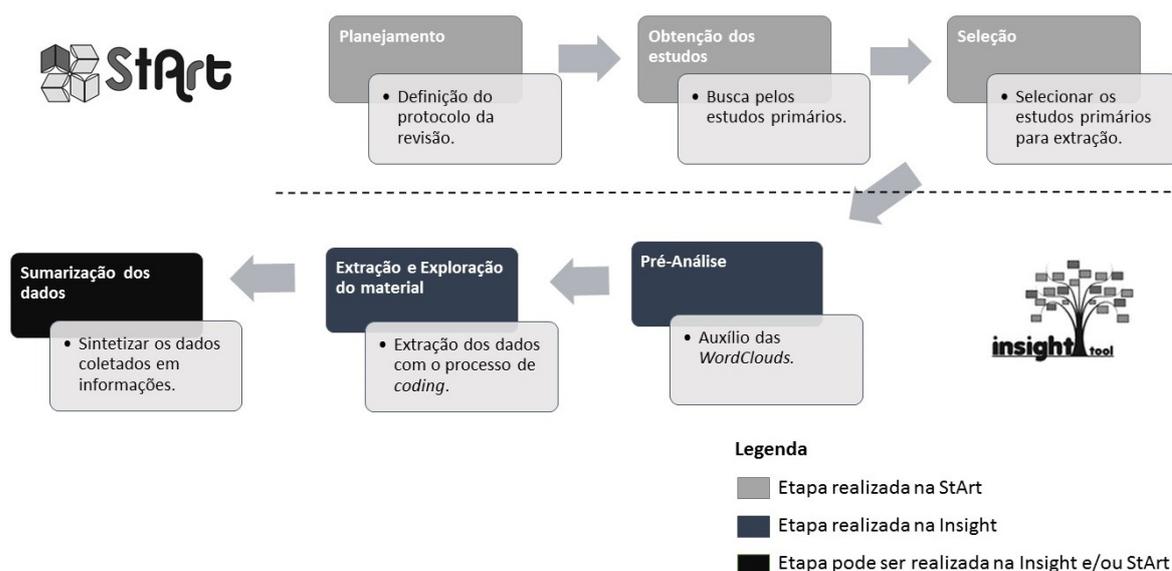
Esta seção detalha a abordagem proposta neste trabalho, a abordagem RS/AT, as etapas a serem seguidas para utilizá-la e como as ferramentas StArt e Insight podem auxiliar nesse contexto. Salienta-se que a abordagem está vinculada ao uso dessas ferramentas, pois elas oferecem recursos especiais para facilitar a adoção da análise temática na revisão sistemática.

O planejamento do estudo é realizado na StArt, assim como a seleção do material de estudo a ser analisado. Uma vez selecionados os estudos primários, a extração dos dados e exploração do material é realizado na ferramenta Insight. Por fim, os dados podem ser sumarizados tanto na ferramenta Insight com a elaboração de um sumário narrativo e geração do mapa mental, bem como na StArt, também para elaboração de um sumário narrativo que venha a compor a caracterização do estado da arte do tópico investigado.

A Figura 3.2 apresenta uma visão geral de quais etapas são executadas em cada uma das ferramentas, o fluxo do trabalho e uma descrição das tarefas realizadas e recursos utilizados.

Uma descrição elaborada de cada etapa ilustrada anteriormente é apresentada a seguir:

1. *Planejamento*: Assim como na revisão sistemática tradicional, ao iniciar o estudo, o primeiro passo que deve ser realizado é o planejamento. Definir o tema a ser tratado, objetivos a serem atingidos, escopo da pesquisa, dados a serem coletados e outras informações



**Figura 3.2: Análise temática em revisão sistemática apoiada por suporte computacional**

que sejam relevantes no contexto.

No planejamento da revisão sistemática tradicional é definido o formulário de extração de dados, os quais correspondem às informações relevantes que se deseja extrair dos estudos primários. Já com a análise temática não é necessária a definição do formulário, porém, deve-se dar atenção especial à elaboração de guias para extração de dados (por exemplo as categorias semente citadas no Capítulo 2) quando há mais de uma pessoa envolvida na análise.

Um dos princípios da análise qualitativa e, conseqüentemente, da análise temática é o caráter exploratório da pesquisa, no qual não se limita os dados que podem ser extraídos. A criação das categorias semente e demais diretrizes que possam ser estabelecidas não impedem que, durante a pesquisa, sejam criadas novas categorias e dados que não correspondem ao contexto daquelas categorias sejam coletados. Apenas fornece alguns parâmetros de dados que devem ser extraídos caso presentes, mantendo assim o caráter exploratório da pesquisa.

- Obtenção dos estudos:* Seja o estudo secundário conduzido com ou sem o apoio da análise temática, o material de estudo deve ser adquirido. Para tal podem ser utilizadas as bases indexadoras de estudos (IEEE, ACM, Scopus, entre outras), indicação de especialista, busca manual, entre outros modos.
- Seleção:* A seleção dos estudos coletados, assim como a obtenção dos estudos, também é comum aos processos com e sem a aplicação da análise temática. A partir dos estudos

The screenshot shows the 'Protocol' planning stage in the StArt software. The interface is divided into a sidebar on the left and a main form area. The sidebar contains navigation options: 'Planning', 'Execution', and 'Summarization'. The main form area is titled 'Protocol' and contains several sections:

- Objective:** A text box containing 'Identificar técnicas de visualização de dados para utilizar em análise qualitativa de dados.'
- Main question:** A text box containing 'Quais técnicas de visualização são utilizadas para a análise qualitativa?'
- Population:** A text box containing 'Estudos da área de Engenharia de Software'
- Intervention:** A text box containing 'Técnicas e aplicações no contexto.'
- Control:** A text box containing 'Fernandes (2014).'
- Results:** A text box containing 'Técnicas de visualização de dados.'
- Application:** A text box containing 'Pesquisadores que desejam identificar técnicas de visualização em análise qualitativa e sua utilização.'
- Keywords and Synonyms:** A section with a list of keywords: 'análise qualitativa', 'dados de visualização', 'informação retrieval', 'qualitative analysis', and 'visualização de dados'. There are buttons for 'Add', 'Remove', 'Up', and 'Down'.

**Figura 3.3: Planejamento da revisão na StArt**

realizados para este trabalho, verificamos que a contribuição para da análise temática nesta etapa é pequena ou inexistente.

Estas conclusões foram obtidas a partir de uma revisão sistemática realizada sobre automatização do processo de seleção em revisões sistemáticas. Constatamos que existem diversas pesquisas que visam automatizar a etapa de seleção dos estudos, ou parte dela, com utilização de diversos mecanismos computacionais. Como pesquisas no enfoque de automatização da etapa de seleção vale citar os estudos de Felizardo et al. (2011) e Octaviano et al. (2014).

4. *Pré-análise:* A etapa de pré-análise é tratada no contexto da análise temática. Consiste de, a partir dos estudos selecionados, efetuar uma leitura flutuante nos documentos a fim de identificar conceitos que possam ser recorrentes e dados que se correlacionem.

O auxílio da visualização implementada na Insight contribui para esta etapa. Utilizando as *Word Clouds* (Figura 3.4), os temas recorrentes no texto são destacados visualmente para o usuário, possibilitando assim obter o contexto da pesquisa e tópicos principais tratados naquele estudo.

Ressalta-se que a pré-análise ainda não é a fase de extração de dados em si, mas sim a imersão no conteúdo dos documentos para ter a visão do todo (conjunto de documentos) e, após, analisar a parte (documentos isolados).

5. *Extração/Exploração do material:* Esta é a etapa de extração dos dados em si. O material de interesse já foi devidamente selecionado e o pesquisador já tem conhecimento dos tópicos tratados nos documentos.

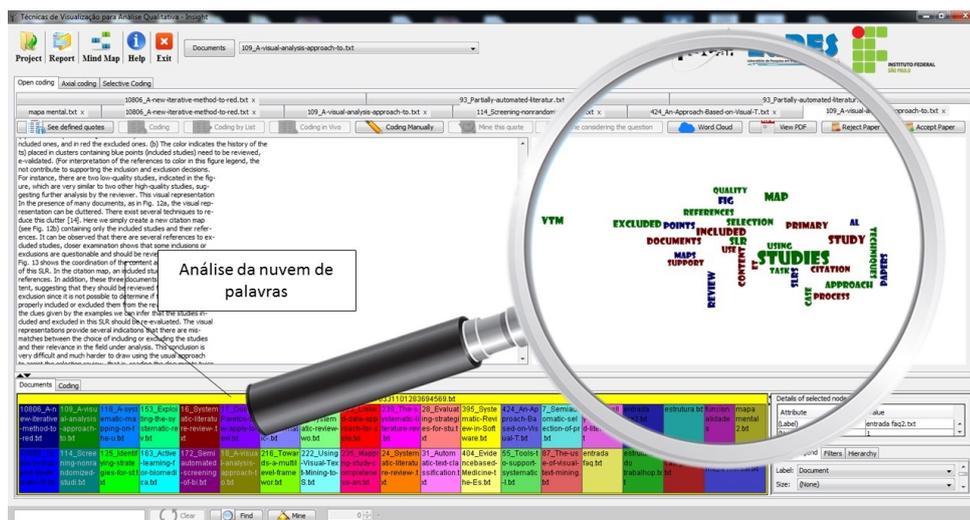


Figura 3.4: Nuvem de palavras

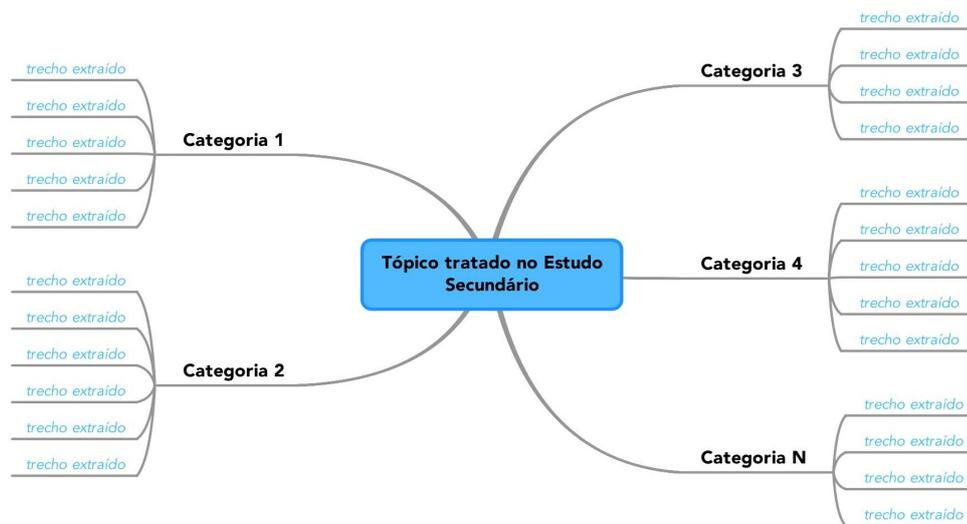
Esta etapa pode ser considerada uma das mais trabalhosas da revisão. Deve-se ler todos os estudos primários e extrair os dados de interesse para a sumarização.

Aplicando a técnica de coding, tratada na seção 2 deste trabalho, o pesquisador irá extrair os trechos de interesse e organizá-los em categorias. Nesta etapa deve-se verificar se foram definidas as mencionadas diretrizes no planejamento da revisão para não omitir nenhuma evidência na extração.

6. *Sumarização/tratamento dos resultados e interpretação*: Por fim, deve-se sumarizar todos os dados extraídos. Visto que os dados estão organizados em categorias, além do sumário narrativo a ser realizado, diferentes formas de visualização podem ser geradas. Uma delas é a criação de um mapa mental (Figura 3.5) dos dados extraídos e categorias geradas. A ferramenta Insight possibilita exportar um arquivo texto já formatado para inserção em software de visualização de mapas mentais, por exemplo o Mindmeister.

Os passos descritos anteriormente auxiliam o usuário com uma abordagem proposta para utilização de análise temática em revisão sistemática. Além da abordagem proposta, neste trabalho investigamos como o suporte computacional pode auxiliar na condução da mesma.

Assim como a abordagem proposta neste trabalho, existem outras pesquisas na área de Engenharia de Software e em diversas áreas do conhecimento que buscam investigar a aplicação de técnicas de análise qualitativa em revisões sistemáticas e propor a aplicação das mesmas. Entre elas, pode-se citar os trabalhos de Thomas e Harden (2008) que propõe a síntese temática. A síntese temática baseia-se nos princípios da análise temática e é aplicada no contexto de revisões sistemáticas.



**Figura 3.5: Modelo de mapa mental**

Os estudos de Cruzes e Dyba (2011) buscam avaliar a utilização e contribuição da síntese temática na área de Engenharia de Software. As pesquisas citadas mostram que o tópico exige extensa discussão devido ao caráter subjetivo da pesquisa qualitativa e são importantes para a área.

Além dos citados estudos, verificamos que o desenvolvimento de pesquisas na área de análise de dados provenientes de estudos primários em Engenharia de Software tem se intensificado. A utilização de abordagens qualitativas torna-se importante para a área de Engenharia de Software e propostas para utilização e sistematização das técnicas, assim como propõe os trabalhos citados e a abordagem RS/AT, podem ser utilizadas como base para aplicação da análise qualitativa em revisões sistemáticas.

### 3.3 Integração e melhorias das ferramentas

As duas seções anteriores apresentaram a caracterização da análise temática no contexto de estudos secundários e a proposição de uma abordagem para aplicação da mesma com auxílio das ferramentas StArt e Insight. Ambas foram desenvolvidas com propósitos específicos - StArt para auxiliar a condução de estudos secundários e Insight para análise de dados qualitativos.

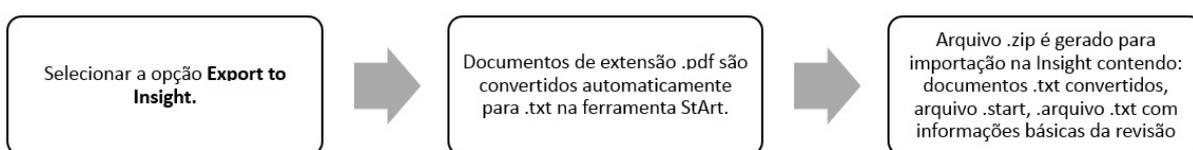
Visto que a proposta deste trabalho é a utilização dos processos em conjunto e, como um dos objetivos a serem alcançados é fornecer um suporte computacional às atividades, foram adicionadas às ferramentas funcionalidades e o compartilhamento de dados entre as mesmas para integração. A seguir são descritos principais desenvolvimentos realizados.

- **Integração das ferramentas StArt e Insight**

Um desenvolvimento fundamental a ser realizado para atingir um dos objetivos do trabalho foi a integração entre as ferramentas. Esta foi realizada após o estudo sobre os processos de revisão sistemática e análise temática em conjunto pois, somente a partir deste, foi possível identificar em quais etapas haveria a possibilidade de integração. Foi constatado que a análise temática pode agregar ao processo de revisão sistemática na etapa de extração e sumarização dos dados - visto que é uma técnica para análise de dados. A mencionada integração das ferramentas foi realizada com base em troca de arquivos entre as mesmas. O fluxo simplificado dos dados é descrito a seguir:

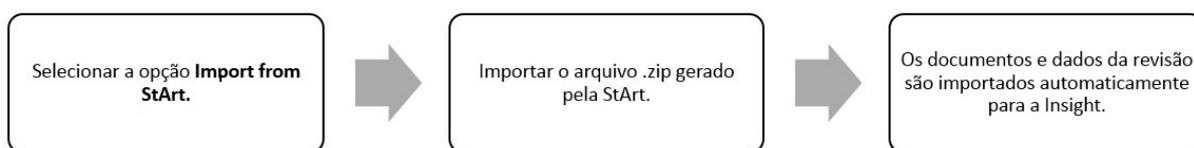
1. *Exportação dos dados da Start*: A exportação dos dados da StArt para trabalho na ferramenta Insight é realizado após selecionados os estudos primários. De posse do material devidamente selecionado, é permitido realizar a exportação utilizando a opção "Export to Insight".

Na ferramenta StArt, os estudos primários completos de extensão .pdf são indexados aos correspondentes arquivos do item na revisão. Atualmente, a ferramenta Insight suporta arquivos de extensão .txt e .xls. A fim de possibilitar a portabilidade dos arquivos, os documentos .pdf associados à StArt são convertidos pela ferramenta para documentos de extensão .txt. Após a conversão, estes arquivos convertidos e os documentos relativos a revisão são comprimidos para outro arquivo de extensão .zip.



**Figura 3.6: Fluxo dos dados para exportação da StArt**

2. *Importação dos dados na Insight*: Para efetuar a análise dos estudos primários selecionados com a StArt, basta abrir a ferramenta Insight e selecionar a opção "Import from StArt". Deve-se então selecionar o arquivo de extensão .zip exportado pela StArt. O arquivo .script (relativo ao projeto da ferramenta Insight) é criado com as informações da revisão e todos os estudos primários da mesma são importados na ferramenta.
3. *Análise dos dados com a Insight*: Uma vez importado o material de análise, cabe ao pesquisador efetuar a análise dos dados com os recursos disponíveis na ferramenta.



**Figura 3.7: Fluxo dos dados para importação na Insight**

Após finalizada a extração dos dados, é possível exportá-los de volta à ferramenta StArt (etapa não obrigatória). Do mesmo modo que é realizada a exportação pela StArt, também é assim realizada pela Insight. É gerado um arquivo de extensão .zip contendo todos os dados relativos a revisão: arquivo .script da Insight, arquivo .txt formatado de estudos primários aceitos e rejeitados, relatório dos dados extraídos, entre outros).

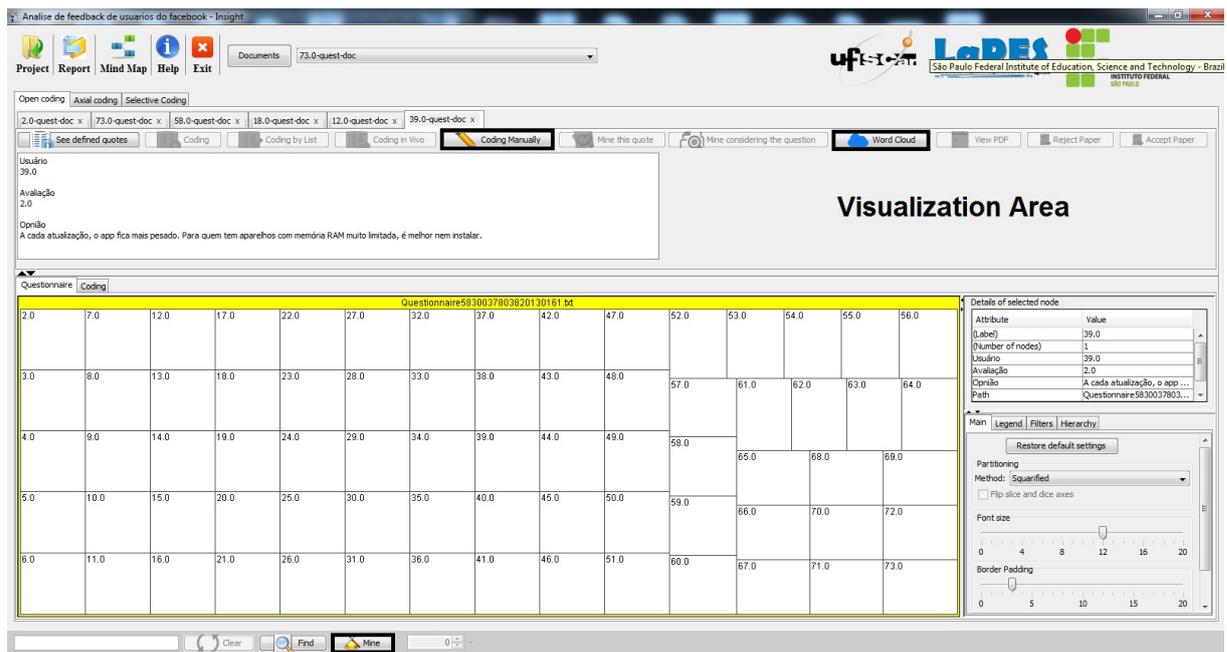
Outras etapas da revisão sistemática foram investigadas a fim de realizar a integração com a técnica de análise temática apoiado pelas ferramentas. Porém, não foram constatados ganhos ao processo e, portanto, a comunicação com as ferramentas é utilizada nas mencionadas etapas. O caminho inverso, ou seja, exportação dos dados analisados da Insight para a StArt, também é possível a fim de utilizar os recursos de visualização de informações disponíveis na StArt - apesar de possível, não é uma etapa obrigatória; o pesquisador pode optar por realizar toda a extração e sumarização na Insight.

Após finalizada a revisão, caso seja necessário atualizá-la, é possível reaproveitar os dados da mesma. Ao realizar a importação dos dados na Insight, verifica-se a existência de arquivo da revisão. Caso sim, é importado este arquivo e os novos documentos ainda não analisados.

Além da integração entre as ferramentas, foram adicionadas algumas funcionalidades na ferramenta Insight para auxiliar à análise dos dados. Entre as funcionalidades incluídas estão a opção de codificação manual dos documentos, mineração de texto baseada em expressão e uma nova opção para visualização de dados (destacados com borda preta na Figura 3.8).

- **Visualização do tipo nuvem de palavras**

Uma das etapas definidas por Bardin (1995) para aplicação da análise temática é a *Pré-Análise*. Consiste de efetuar uma leitura flutuante a fim de obter conhecimento sobre o conjunto de dados a ser analisado. As funcionalidades "Mine" e "WordCloud" auxiliam principalmente esta etapa - Hernandez (2014) destaca em seu trabalho que a visualização de informações contribui para a análise qualitativa de dados. Com o auxílio da visualização e mineração de texto estes recursos implementados auxiliam o pesquisador a identificar tópicos recorrentes, ou não,



**Figura 3.8: Funcionalidades adicionadas na ferramenta para todos os tipos de projeto**

nos dados de entrada. Identificar a similaridade de uma expressão no conjunto de documentos pode auxiliar também na definição das categorias iniciais da análise temática.

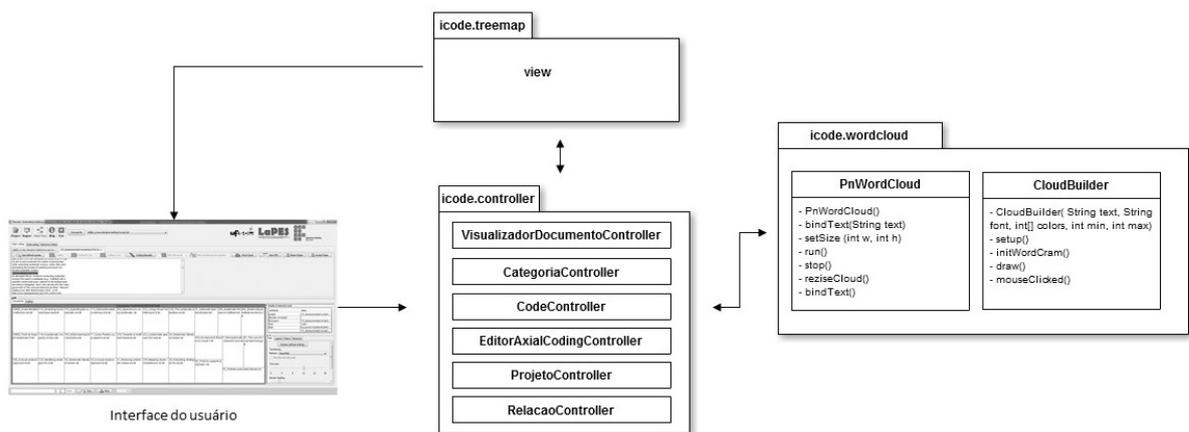
Nesse contexto, a opção de visualização de nuvem de palavras, ou *Word Cloud*, adicionada faz uso da mineração de texto em conjunto com visualização de informações para extrair as palavras recorrentes do texto e disponibilizá-las visualmente ao usuário. A Figura 3.9 apresenta um exemplo de nuvem de palavras gerada pela ferramenta. Este tipo de visualização permite identificar os termos mais recorrentes no documento.



**Figura 3.9: Visualização do tipo WordCloud gerada pela Insight**

A implementação da nuvem de palavras foi realizada com auxílio de uma biblioteca aberta, a WordCram. Do ponto de vista da estrutura de funcionamento do algoritmo, é realizado de acordo com o esquema ilustrado na Figura 3.10 e descrito nas etapas a seguir:

1. Requisição do usuário: A partir do clique do mouse no botão "WordCloud", é iniciado a chamada do método para geração da nuvem de palavras.
2. Aquisição do conteúdo: O controlador do sistema obtém o conteúdo do documento corrente e envia a requisição para as classes responsáveis pela nuvem de palavras.
3. Cálculo de palavras recorrentes: As palavras do documento são ranqueadas de acordo com a incidência no texto.
4. Visualização do componente: De acordo com as configurações estabelecidas na implementação, são disponibilizadas ao usuário as 20 palavras mais recorrentes no texto. Estas serão disponibilizadas em tons de cores e com fonte pré-determinada em um painel da interface do usuário.



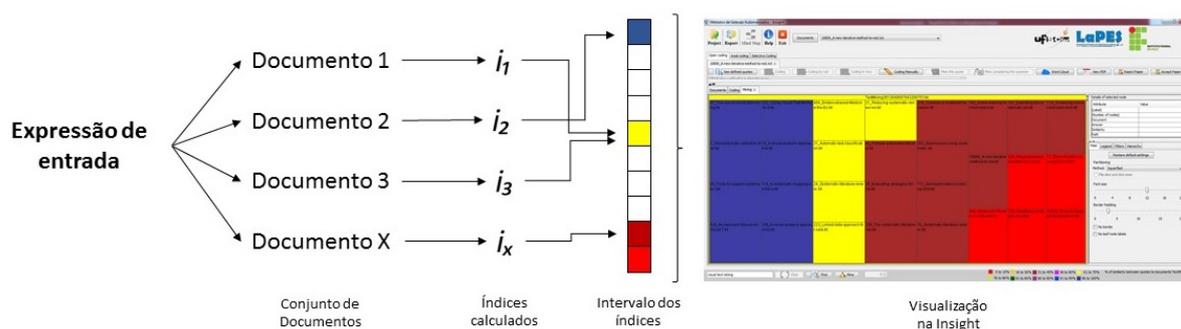
**Figura 3.10: Estrutura de funcionamento da WordCloud**

O componente ainda possui algumas limitações que devem ser aprimoradas. A principal trata-se de composição de expressões (Ex: Engenharia de Software, Revisão Sistemática, entre outras). Atualmente, as expressões são tratadas como palavras separadas para se efetuar o ranqueamento na nuvem de palavras - "Engenharia de Software" é uma expressão composta por três palavras, atualmente cada palavra é ranqueada separadamente. Como trabalho futuro, esperamos aprimorar este componente para efetuar o ranqueamento de acordo com expressões pré-estabelecidas em um dicionário de domínio do sistema.

- **Mineração de texto baseada em expressão**

Outra funcionalidade adicionada e que faz uso de alguns recursos de mineração de texto já presentes na ferramenta foi o botão "*Mine*".

Utiliza como base o algoritmo da janela deslizante, mencionado na seção sobre a Insight, tendo como entrada uma expressão definida pelo usuário para cálculo da similaridade em cada documento. A Figura 3.11 ilustra a arquitetura do código para efetuar o cálculo. O usuário insere uma expressão de entrada e, para cada documento, é calculado um índice utilizando o algoritmo da janela deslizante. Estes índices são valores numéricos que, após calculados, são alocados em uma escala composta por 10 intervalos. Cada um destes intervalos será representado por uma cor e cada documento terá atribuído a si a cor do intervalo correspondente na visualização disponibilizada pela Insight.



**Figura 3.11: Estrutura de funcionamento do botão Mine**

- **Codificação manual**

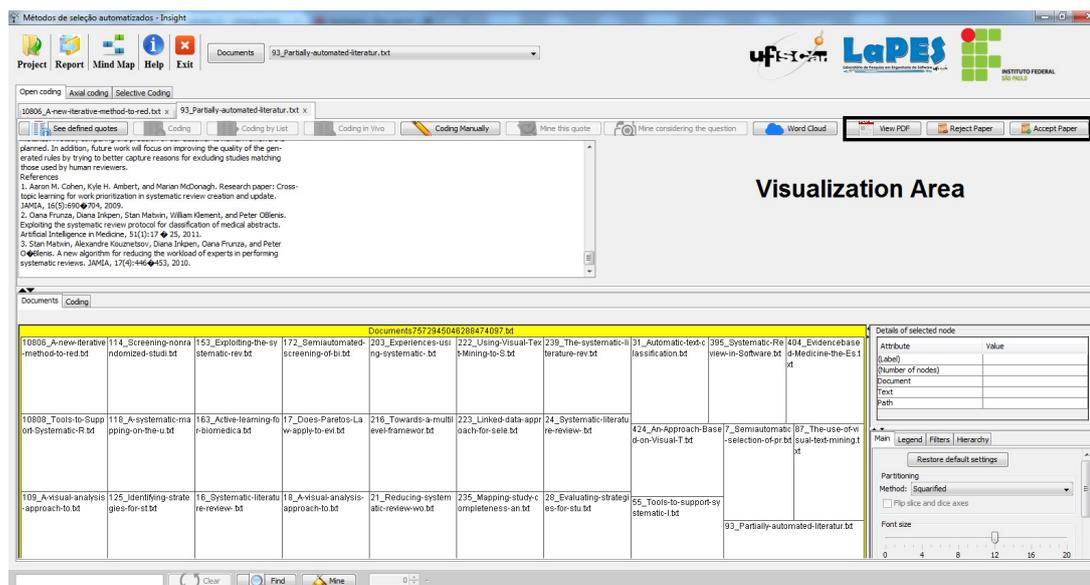
A codificação manual, botão intitulado "*Coding Manually*", é mais uma opção para codificação dos documentos disponibilizado na ferramenta. Permite que o usuário insira tanto o rótulo (*code*) quanto a passagem para rotulação (*quotation*) manualmente.

Esta opção pode ser utilizada também para inserir notas ou anotações a respeito do documento analisado.

- **Opções específicas para condução de estudo secundário**

Além das mencionadas funcionalidades, foram adicionadas à ferramenta opções exclusivas para condução de estudos secundários (destacados em borda preta na Figura 3.12, botões "*View PDF*", "*Accept Paper*" e "*Reject Paper*").

A opção "*View PDF*" permite a visualização do arquivo PDF associado ao documento corrente no visualizador padrão de PDF do sistema do usuário, a fim de permitir ao pesquisador efetuar anotações nos documentos originais também. No clique do botão, a ferramenta lança uma chamada ao sistema para abertura do documento associado com o visualizador padrão instalado. Caso não haja um visualizador instalado, é recomendado ao usuário fazê-lo.



**Figura 3.12: Funcionalidades exclusivas para projetos advindos da ferramenta StArt**

Já as opções "*Accept Paper*" e "*Reject Paper*" funcionam de modo semelhante. Ao clicar em alguma das opções, é inserido no documento corrente a situação do mesmo na revisão, ou seja, se foi aceito ou rejeitado nesta etapa. Como mencionado na Seção 2, mesmo durante a extração dos dados em estudos secundários, o pesquisador tem a possibilidade de aceitar e rejeitar um estudo primário após a leitura completa do documento. Estas informações são disponibilizadas ao encerrar a revisão para documentação da mesma.

### 3.4 Considerações finais

Este capítulo apresentou os suportes computacionais utilizados para este trabalho bem como os desenvolvimentos realizados para possibilitar a aplicação da análise temática em conjunto com a revisão sistemática.

Como já mencionado anteriormente, o principal aspecto da análise qualitativa, conjunto maior no qual a análise temática está contida, é o caráter interpretativo deste tipo de pesquisa. Neste contexto, as técnicas computacionais disponibilizadas devem fornecer mecanismos para auxiliar o pesquisador na interpretação dos dados - motivação pela qual as técnicas de mineração de texto e visualização de dados foram implementadas na Insight.

# Capítulo 4

## ESTUDOS CONDUZIDOS

---

---

### 4.1 Considerações iniciais

Considerando a abordagem de aplicação da análise temática na condução da revisão sistemática, apresentada no capítulo anterior, este capítulo apresenta alguns estudos que foram realizados com o objetivo de avaliar a abordagem.

Foram realizados três estudos. O primeiro teve como objetivo levantar dados de pesquisadores que conduzem revisão sistemática e/ou análise qualitativa de dados quanto a utilização e grau de conhecimento nas mencionadas técnicas. O segundo estudo explorou a comparação da abordagem desenvolvida neste trabalho (análise temática aplicada no contexto de revisão sistemática) em relação a abordagem tradicional utilizada de revisão sistemática (utilizando formulários de extração). Finalmente, o terceiro estudo também efetuou uma comparação entre as abordagens mencionadas a fim de validar hipóteses estabelecidas no segundo estudo.

Todos os estudos foram planejados seguindo as diretrizes propostas por Wohlin et al. (2000), sendo que na definição dos estudos adotou-se o modelo GQM - Goal, Question, Metric (BASILI; ROMBACH, 1988).

Com esses três estudos procurou-se explorar a abordagem proposta de diferentes pontos de vista, desde a utilização por potenciais usuários, passando por pesquisadores iniciantes e explorando até pesquisadores experientes em revisão sistemática.

A análise dos dados decorrentes dos estudos realizados, assim como parte do tema deste trabalho, foi realizada utilizando análise qualitativa.

Este capítulo está organizado conforme descrito a seguir. A seção 4.2 detalha o *survey* aplicado a fim de identificar potenciais usuários para a abordagem RS/AT, desenvolvida neste

trabalho. Na seção 4.3 é detalhado o estudo piloto realizado comparando a abordagem RS/AT com a revisão sistemática utilizando formulário de extração. Finalmente, a seção 4.4, detalha o experimento realizado para validar as hipóteses estabelecidas no estudo piloto anterior, descrito na seção 4.3, quanto a contribuições e limitações da utilização da abordagem RS/AT em relação a revisão sistemática com formulário de extração para extração dos dados.

## 4.2 Survey para avaliação de perfil de usuários

Esta seção apresenta o *survey* aplicado para levantamento de possíveis usuários interessados em utilizar revisão sistemática e/ou análise qualitativa. Nas subseções a seguir são detalhados a definição, planejamento, operação, análise e interpretação dos dados e ameaças à validade do estudo conduzido.

### 4.2.1 Definição

A fim de identificar potenciais usuários para a abordagem RS/AT e qual apoio ferramental poderia ser utilizado, foi disponibilizado um *survey* online que procurou atingir, principalmente, pessoas que já tiveram contato com revisão sistemática e/ou análise qualitativa.

Este survey foi conduzido na etapa inicial da pesquisa, antes de ser definida a abordagem RS/AT em si e a implementação da integração das ferramentas StArt e Insight. Foi utilizado como um ponto de partida para avaliar possíveis usuários de uma abordagem que integrasse revisão sistemática e análise temática com utilização de ferramentas.

Vários fatores são responsáveis pela resistência à utilização de novas técnicas, ou mesmo de técnicas já conhecidas: não conhecimento da abordagem ou técnica, dificuldade de aplicação da abordagem/técnica, falta de ferramentas que auxiliem a condução da abordagem/técnica, entre outros.

Neste contexto, este estudo avaliou o perfil dos respondentes sendo que, com base nesse perfil, identifica-se se haveria pessoas que teriam conhecimento nas técnicas envolvidas, de modo a se tornarem potenciais usuários da abordagem RS/AT.

### 4.2.2 Planejamento

O planejamento deste estudo foi conduzido de acordo com o modelo GQM (BASILI; SELBY; HUTCHENS, 1986) e é apresentado na Tabela 4.1.

Para coletar os dados do estudo foi definido um formulário *online* (disponível no Apêndice A) disponibilizado em lista de emails e redes sociais, de modo anônimo. Ressalta-se que os alvos estabelecidos foram: a lista de pós-graduandos do PPGCC/Ufscar e a lista de alunos da UFSCar (com solicitação de encaminhamento para outros interessados), uma vez que nesses fóruns era de conhecimento que várias pessoas possuem conhecimento das técnicas investigadas.

**Tabela 4.1: Definição do *survey* para avaliação de potenciais usuários**

<b>Objeto analisado</b>	O perfil dos respondentes do <i>survey</i>
<b>Com o propósito de</b>	Identificar potenciais usuários da abordagem RS/AT
<b>Com respeito</b>	ao grau de conhecimento em RS; ao procedimento usado para extração dos dados em RS; ao grau de conhecimento em análise qualitativa; a qual técnica de análise qualitativa é utilizada
<b>Do ponto de vista dos</b>	Pesquisadores
<b>No contexto de</b>	Alunos de graduação, pós-graduação e demais pesquisadores

### 4.2.3 Operação

A seguir são apresentadas as questões e as respostas dos 21 participantes do *survey*.

*Questão 1 - Qual destes perfis se adequa ao seu?*

A primeira questão identificou o perfil dos respondentes. O perfil da maioria dos respondentes foi de estudantes de pós-graduação, com total de 14 indivíduos (entre níveis mestrado e doutorado) - quatro estudantes de graduação, dois mestres e um graduado, conforme mostra a Figura 4.1.

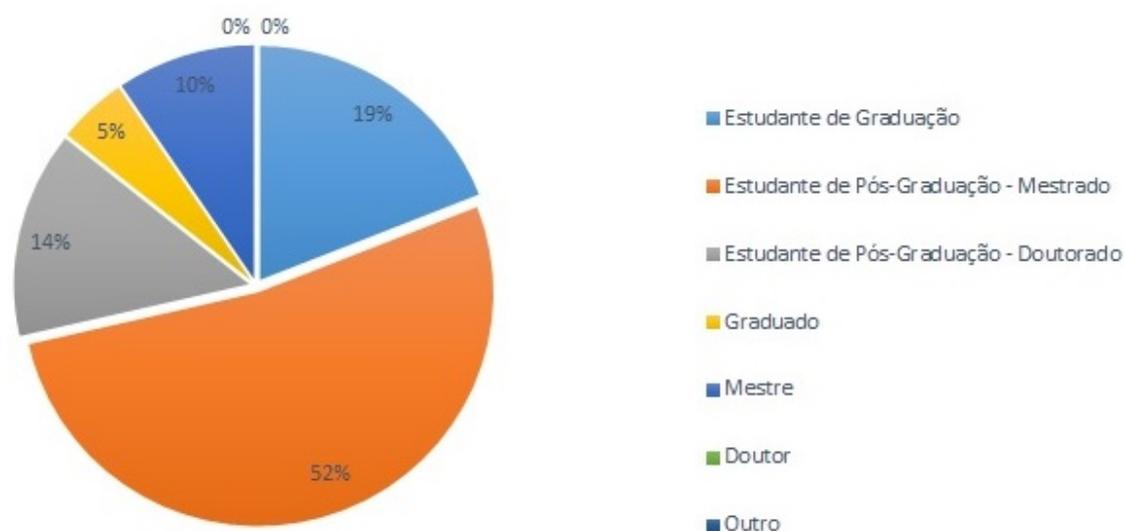
*Questão 2 - Qual seu curso?*

No que diz respeito ao curso dos respondentes, o curso de Ciência da Computação foi o que se destacou com maior número de participantes. Este resultado certamente foi decorrente do fato de que o *survey* foi disponibilizado na lista de estudantes do PPGCC/UFSCar (Programa de Pós Graduação em Ciência da Computação da UFSCar).

Em seguida, o curso com mais respondentes foi o de Educação Especial. As técnicas de análise qualitativa de dados e revisão sistemática da literatura são muito utilizadas na área de ciências humanas e biológicas.

*Questão 3 - Conduz ou já conduziu estudos secundários? Se sim, quais tipos?*

Esta questão permitia ao respondente a escolha de mais de uma opção a fim de identificar quais tipos de estudos secundários eram de conhecimento dos respondentes. Neste conjunto de



**Figura 4.1: Respostas Q1 - Qual destes perfis se adequa ao seu?**

participantes predominou a revisão sistemática. Treze participantes responderam que conduzem ou já conduziram uma revisão sistemática, quatro conduzem revisão ad-hoc, seis nunca conduziram qualquer tipo de estudo secundário e um participante citou que já conduziu ou conduz mapeamento sistemático (listado como "outro").

A maior quantidade de respostas para revisão sistemática foi de participantes da área de ciência da computação.

#### *Questão 4 - Qual seu grau de experiência em estudos secundários?*

A fim de avaliar e ponderar as respostas da Questão 3, o grau de experiência dos usuários foi questionado em uma escala de 1 a 5, sendo 1 como "sem experiência" e 5 como "experiente". As respostas coletadas são apresentadas na Figura 4.4.

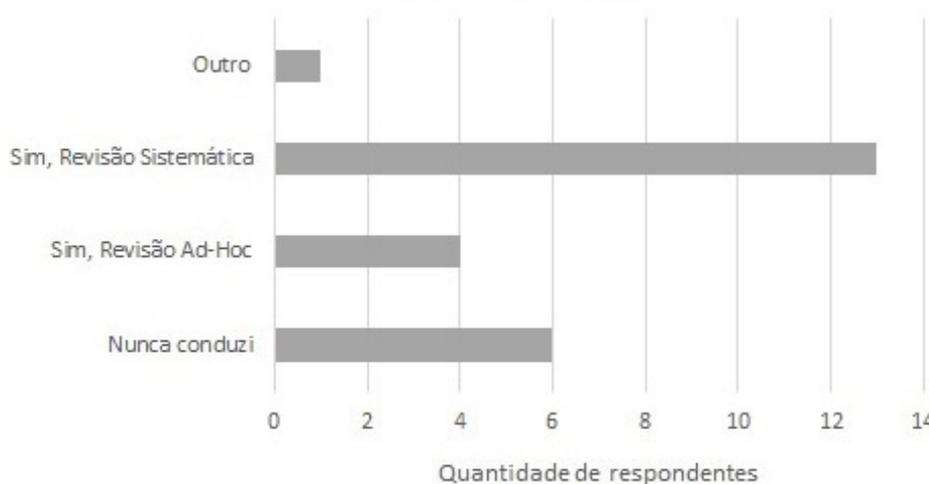
#### *Questão 5 - Como efetua a análise dos estudos primários?*

Aos participantes que conduzem estudos secundários, foi solicitado que descrevessem como efetua a análise dos estudos primários.

O modo mais citado de realizar a extração dos dados foi a utilização de formulários de extração (15 participantes). Esse número é condizente com as respostas da Questão 3, na qual 15 participantes responderam que conduzem revisões sistemáticas - formulários de extração são o meio indicado para se efetuar a extração dos dados de estudos primários nesse tipo de estudo secundário, ou seja, os 15 participantes que responderam que efetua revisões sistemáticas utilizam os formulários de extração.



**Figura 4.2: Respostas Q2 - Qual seu curso?**

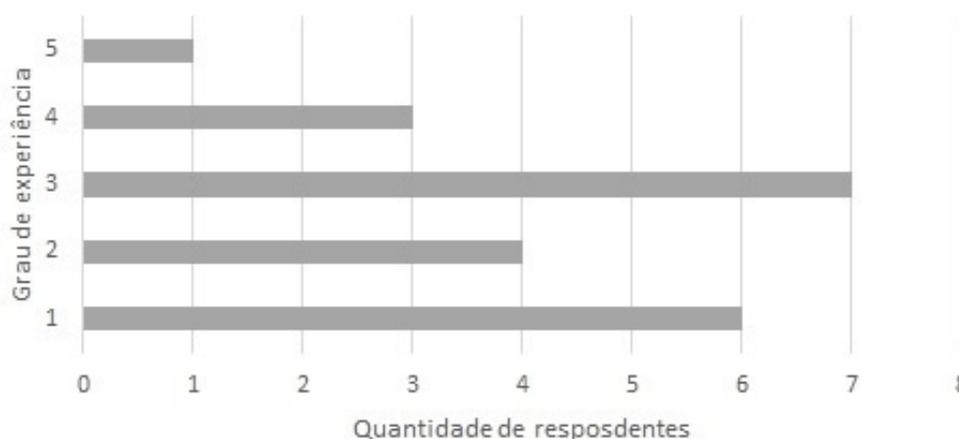


**Figura 4.3: Respostas Q3 - Conduz ou já conduziu estudos secundários? Se sim, quais tipos?**

Outro modo de análise dos estudos primários, citado por 5 dos respondentes, foi a elaboração de resumos dos estudos primários para posterior recuperação e análise dos mesmos em conjunto a fim de elaborar as conclusões. Observa-se que, nesse caso, um modo que poderia ser mais eficiente nestes casos, visto que optam por não estabelecer formulários, seria a aplicação da análise temática em si para extração dos dados ou alguma outra técnica de análise qualitativa.

As outras respostas obtidas para esta questão não foram descritas pois não tratavam de análise dos estudos primários ou não descreviam de modo claro como a análise foi conduzida.

*Questão 6 - Considera relevante o desenvolvimento de novas técnicas para auxiliar a extração e sumarização dos dados de estudos secundários?*



**Figura 4.4: Respostas Q4 - Qual seu grau de experiência em estudos secundários?**

Para esta questão, 18 participantes responderam que consideram relevante o desenvolvimento de novas técnicas para auxiliar a condução de estudos secundários e 3 que não consideram relevante.

*Questão 7 - Conduz ou já conduziu análise qualitativa de dados? Se sim, utilizou alguma técnica?*

A fim de identificar o grau de conhecimento dos respondentes em técnicas de análise qualitativa, foi efetuada a questão 7. As respostas são ilustradas na Figura 4.5.

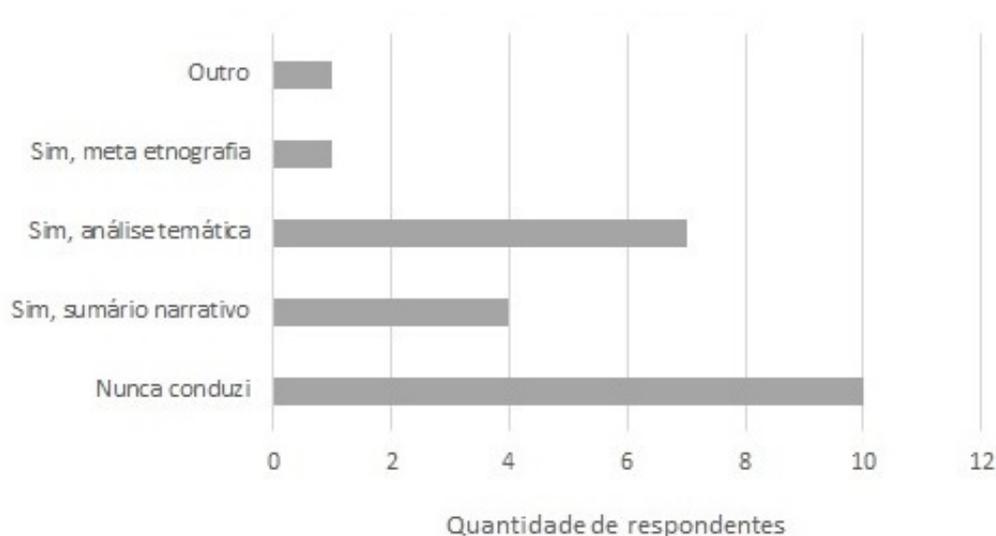
Principalmente na área de ciências exatas, as abordagens qualitativas ainda não são de uso frequente - panorama que vem sendo alterado há algum tempo, principalmente no campo de Engenharia de Software, visto que a demanda por novos mecanismos e abordagens de análise de dados tem crescido (TRAVASSOS, 2002).

A maior parte dos participantes, 10 deles, responderam nunca ter conduzido análise qualitativa de dados. Entre os que conduziram, a técnica mais citada foi a de análise temática. Além da análise temática, outras técnicas citadas foram o sumário narrativo e um dos participantes citou ter utilizado meta-etnografia.

*Questão 8 - Qual seu grau de experiência em análise qualitativa de dados?*

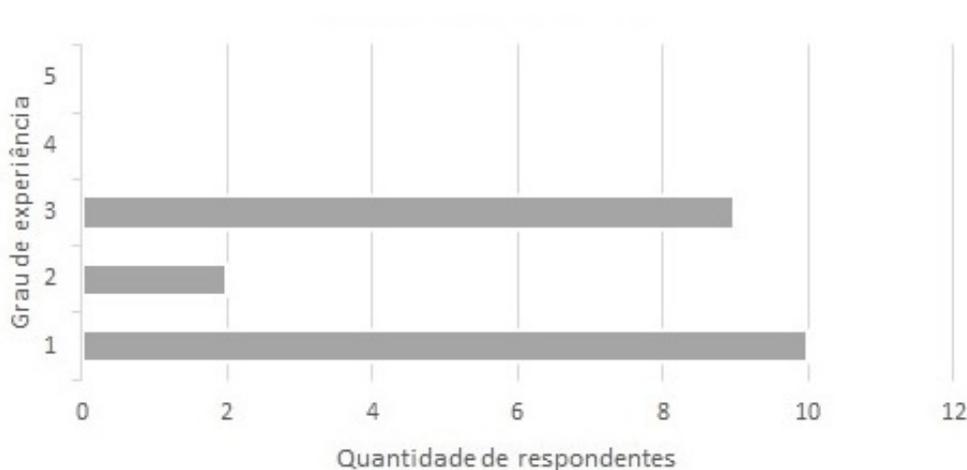
De modo análogo à pergunta sobre o grau de experiência dos participantes em estudos secundários, foi perguntado sobre o grau de experiência em análise qualitativa de dados, também obedecendo uma escala de 1 (menor) a 5 (maior).

Ponto a ser destacado é que nenhum dos participantes assinalou possuir experiência em análise qualitativa de dados maior que 3, na escala disponibilizada. Tal fato pode ser decorrente



**Figura 4.5: Respostas Q7 - Conduz ou já conduziu análise qualitativa de dados? Se sim, utilizou alguma técnica?**

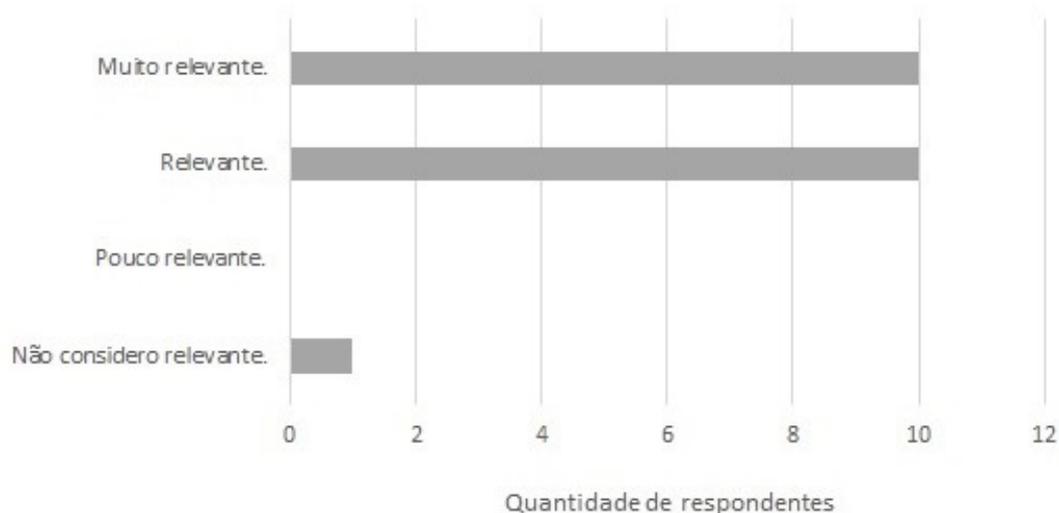
do perfil dos respondentes, uma vez que a maioria é da área de exatas e, aparentemente, tendem a utilizar abordagens quantitativas.



**Figura 4.6: Respostas Q8 - Qual seu grau de experiência em análise qualitativa de dados?**

*Questão 9 - Considera relevante a aplicação de análise qualitativa em estudos secundários?*

A última questão foi relativa à opinião do usuário sobre a relevância do objetivo principal deste trabalho: aplicação de análise qualitativa (no caso, a análise temática) em estudos secundários. Das respostas coletadas, 20 consideraram como "Muito relevante" ou "Relevante" e 1 considerou como "Não considero relevante" (Figura 4.7).



**Figura 4.7:** Respostas Q9 - Considera relevante a aplicação de análise qualitativa em estudos secundários?

#### 4.2.4 Análise e interpretação

Ao analisar as respostas deste *survey*, obtidas na Questão 2, verificou-se que grande parte dos participantes são da área de ciência da computação. Esse resultado é decorrente pelo meio de divulgação do *survey*, como mencionado anteriormente.

Muitos dos participantes já conduziram um estudo secundário. A maior parte citou na Questão 3 a utilização da revisão sistemática e, conseqüentemente a utilização de formulários de extração para sumarização dos dados. Ao analisar a Questão 4 quanto ao grau de experiência dos participantes em estudos secundários, mais da metade julga possuir experiência média ou menor.

Após abordado o conceito de grau de conhecimento sobre estudos secundários, foi questionado aos participantes o grau de conhecimento em análise qualitativa de dados. Verificou-se com as respostas das questões 7 e 8 que poucos dos participantes tem conhecimento de alguma técnica de análise qualitativa de dados e, quando possuem, o grau de experiência na utilização da análise qualitativa é mediano ou menor. Tal fator pode ser associado ao perfil dos participantes - maior parte de exatas, área na qual a análise quantitativa de dados ainda é mais presente. Porém, apesar do pouco grau de conhecimento e aplicação, ao analisar as respostas de texto da questão 7 sobre a técnica utilizada, a técnica mais citada foi a análise temática, seguida por sumário narrativo e meta-etnografia.

Finalmente, as questões 6 e 9 questionaram, respectivamente, sobre o desenvolvimento de novas técnicas para revisão sistemática e sobre a aplicação da análise qualitativa em revisão

sistemática. A maior parte dos participantes considera relevante o desenvolvimento de novas técnicas, entre elas a aplicação de análise qualitativa de dados. Embora as questões 7 e 8 mostrem pouco conhecimento dos respondentes quanto a análise qualitativa de dados, supõe-se que o próprio questionário tenha despertado a curiosidade sobre a análise qualitativa, fazendo com que eles investigassem sua possível aplicação em estudos secundários. Outro fator que pode ter levado a estas respostas é a dificuldade de desenvolver um formulário de extração no início da revisão - o que gera dificuldade para extrair dados na revisão sistemática que podem ser a adição/remoção de itens do formulário até a necessidade de analisar novamente os estudos.

Baseado no panorama apresentado sobre potenciais usuários interessados em revisão sistemática ou análise qualitativa dos dados, julgamos que a aplicação da análise temática como método de extração e sumarização dos dados pode trazer contribuições para a área e devem ser objeto de análise detalhada. O apoio computacional de ferramentas, explorado também neste trabalho, contribui para este propósito pois a quantidade de dados com o qual são trabalhados em revisão sistemática é grande e conduzir análise qualitativa é trabalhoso. Em vista disso, foi desenvolvida a abordagem proposta neste trabalho e detalhada no Capítulo 3. Os estudos de caso e experimento, listados a seguir, investigaram a aplicação da abordagem proposta em comparação ao processo tradicional da revisão sistemática, utilizando formulários de extração.

#### **4.2.5 Ameaças à validade**

Como ameaças a validade deste estudo podemos organizá-las em validade interna, externa, de construção e de conclusão.

Quanto a validade interna do estudo, referentes a fatores não controlados que podem influenciar as variáveis do estudo, citamos:

- **Participantes:** O *survey* foi disponibilizado a alunos de graduação, pós-graduação e pesquisadores. A resposta dos participantes é livre e válida para aquele contexto de participantes. Os resultados e análises obtidas não podem ser generalizadas para outros contextos de participantes.

Quanto as ameaças externas, referentes a capacidade de generalizar os resultados deste experimento, citamos:

- **Homogeneidade dos participantes:** Muitas das respostas obtidas foram de participantes do PPGCC-UFSCar. Os alunos do programa possuem uma matéria específica de revisão

sistemática. Deste modo, algumas respostas podem ser baseadas nessa experiência prévia. Em vista disso, ressalta-se que a análise e conclusão obtida são válidas para este contexto de participantes.

- Interpretação das questões pelos participantes: O *survey* foi disponibilizado online aos participantes. Eventuais dúvidas quanto as questões disponibilizadas não foram levantadas. Porém, ressalta-se que o e-mail do pesquisador responsável pelo *survey* foi disponibilizado caso ocorresse alguma dúvida ao responder o questionário.

Quanto as ameaças de conclusão, relacionadas a interpretação dos dados, citamos:

- Interpretação das respostas dos participantes: A maior parte das questões era de cunho quantitativo e não sujeito a interpretação. Porém algumas questões eram interpretativas e foram analisadas e interpretadas para obter uma análise mais detalhada - a interpretação foi realizada pelo pesquisador e minimizada, porém ainda é uma ameaça à validade.

Quanto a ameaça de construção do estudo, relacionada às métricas utilizadas para avaliar o estudo, citamos:

- Métricas utilizadas: Algumas das métricas utilizadas como grau de experiência em revisão sistemática ou grau de experiência em análise qualitativa, obtidas das questões do *survey*, foram obtidas a partir da avaliação pessoal do participante quanto a própria experiência. Essas métricas foram utilizadas apenas com intuito de caracterizar os participantes e ponderar possíveis observações deles sobre as técnicas mencionadas.

### 4.3 **Estudo Piloto**

A partir da análise dos processos de revisão sistemática e análise temática, foi desenvolvida neste trabalho uma abordagem para aplicação de análise temática em revisão sistemática utilizando suporte computacional.

Esta seção apresenta os detalhes do estudo conduzido com objetivo de comparar os dados extraídos utilizando a abordagem desenvolvida em comparação com a abordagem tradicional e identificar vantagens e dificuldades encontradas com ela em comparação a utilização da abordagem tradicional (com utilização de formulários de extração para extração de dados).

Foi conduzido por três participantes de diferentes níveis quanto aos conhecimentos em análise temática e revisão sistemática. A definição deste estudo é descrita, conforme o GQM (BASILI; SELBY; HUTCHENS, 1986), na Tabela 4.2.

**Tabela 4.2: Definição do Estudo de Caso I**

<b>Objeto analisado</b>	A abordagem RS/AT
<b>Com o propósito de</b>	Avaliar a revisão sistemática com análise temática em comparação com a revisão sistemática utilizando formulário de extração
<b>Com respeito a</b>	Vantagens e desvantagens de sua aplicação, consistência dos dados, dados extraídos e quantidade de trabalho realizado
<b>Do ponto de vista dos</b>	Pesquisadores
<b>No contexto de</b>	Pesquisador iniciante em revisão sistemática e análise temática, especialista em análise temática e especialista em revisão sistemática

### 4.3.1 Seleção do Contexto

Esse estudo piloto foi realizado com um aluno de graduação, um aluno de pós-graduação de nível mestrado (e autor deste trabalho) e um aluno de pós-graduação de nível doutorado, todos alunos de Ciência da Computação da Universidade Federal de São Carlos.

De acordo com Wohlin et al. (2000), o contexto desse experimento pode ser caracterizado, de acordo com as quatro dimensões (processo, participantes, realidade e generalidade), como:

- Processo: o processo é considerado *Off-line* para este experimento pois os participantes não foram entrevistados durante toda a condução do experimento, apenas ao final de cada etapa.
- Participantes: o contexto dos participantes eram estudantes de pós-graduação.
- Realidade: a realidade do experimento é caracterizada como modelada pois as ferramentas utilizadas não foram desenvolvidas pela indústria e o estudo foi executado em ambiente acadêmico.
- Generalidade: a generalidade dos resultados desse experimento é classificada como específica - os resultados do experimento não podem ser generalizados para outros contextos.

### 4.3.2 Seleção das Variáveis

Para esse estudo foi identificada uma variável independente, denominada Técnica de Revisão, a qual representa a abordagem de extração e sumarização de dados utilizada pelos participantes. A variável independente possui dois tratamentos:

- Formulário de extração: representada pela aplicação do formulário de extração para extração dos dados da revisão sistemática, utilizando a ferramenta StArt.
- RS/AT: representada pela aplicação da abordagem RS/AT, proposta neste trabalho, para extração dos dados utilizando a ferramenta Insight.

Quanto as variáveis dependentes, foram identificadas três:

- Omissão de evidências: definido como a comparação dos dados coletados pelos três pesquisadores.
- Retrabalho: definido como a necessidade de analisar novamente estudos que já haviam sido analisados.
- Dados discrepantes: definido como a coleta de dados não pertencentes ao contexto da pesquisa.

### 4.3.3 Projeto Experimental

O projeto experimental deste estudo foi elaborado com três participantes, denominados P1, P2, P3, detalhados na Tabela 4.3. Para avaliar e ponderar os dados do estudo de acordo com a experiência de cada participante com as técnicas utilizadas, foi definida uma escala de 1 a 5 (sendo 1 para pouca experiência com a técnica até 5 para muito experiente).

**Tabela 4.3: Caracterização dos participante do estudo**

Participante	Abordagem utilizada	Experiência em RS	Experiência em AT
<b>P1</b>	RS/AT	5	5
<b>P2</b>	Formulário de extração	5	1
<b>P3</b>	RS/AT	1	1

P1 utilizou a abordagem RS/AT e possui experiência em revisão sistemática e análise temática.

P2 aplicou a abordagem tradicional de condução de revisão sistemática, utilizando formulários de extração, e possui experiência em conduzir revisões sistemáticas.

P3 utilizou a abordagem RS/AT, assim como P1. Não possui experiência em revisão sistemática nem em análise temática. Em decorrência disso, P3 recebeu um treinamento para ensiná-lo os conceitos fundamentais que seriam abordados no estudo.

### 4.3.4 Instrumentação

Os instrumentos utilizados no estudo foram a ferramenta StArt, implementada em Java e utilizada para dar suporte à condução de revisão sistemática, e a ferramenta Insight, também implementada em Java e utilizada para dar suporte à análise qualitativa de dados.

### 4.3.5 Preparação e Execução do Estudo de Caso

A primeira etapa do estudo foi o desenvolvimento do protocolo da revisão sistemática. Este protocolo foi desenvolvido por P2 em conjunto com outros pesquisadores não participantes deste estudo. Depois de elaborado o protocolo, P2 realizou as outras etapas da RS, gerando portanto, os estudos primários selecionados e a extração dos dados, conforme o objetivo proposto no protocolo, dentre outras coisas. Terminada a RS por P2, os dados apresentados na Tabela 4.4 foram disponibilizados para P1 e P3. O protocolo completo da revisão não foi disponibilizado a P1 e P3 pois, deste modo, eles teriam acesso aos campos de extração e demais informações que poderiam guiar sua coleta de dados. Assim, além dos principais itens do protocolo da revisão, foram disponibilizados a P1 e P3 os estudos primários já selecionados por P2, uma vez que o objetivo deste estudo era comparar a etapa de extração dos dados.

**Tabela 4.4: Dados de planejamento da revisão disponibilizados à P1 e P3**

<b>Objetivo</b>	Levantar as estratégias (automáticas ou não) utilizadas para a seleção de estudos primários em revisões sistemáticas.
<b>Questão Principal</b>	Quais estratégias utilizadas para a primeira seleção de estudos primários em Revisão Sistemática?
<b>População</b>	Estudos que apresentem procedimentos ou métodos para relação de estudos primários em Revisão Sistemática.
<b>Intervenção</b>	Métodos e procedimentos para realizar a seleção de estudos primários em Revisão Sistemática.
<b>Controle</b>	Katia Felizardo, Sandra Fabbri, Petersen.
<b>Resultados</b>	Apresentação e discussão sobre os procedimentos e métodos usados na atividade de seleção de estudos primários.
<b>Aplicação</b>	Pesquisadores que conduzem Revisão Sistemática.

Uma vez realizada as atividades descritas, a partir dos estudos primários selecionados, P1 e P3 utilizaram a ferramenta Insight para conduzir a extração dos dados com a abordagem RS/AT. Já P2, utilizou a ferramenta StArt, tendo que definir o formulário de extração de dados.

Para ilustrar a diferença de procedimentos adotados, a Figura 4.8) mostra uma tela da Insight, representando a extração de dados sendo feita por P1 e P3 com a abordagem RS/AT; já a Figura 4.9 mostra uma tela da StArt, representando a extração de dados sendo feita por P2, com

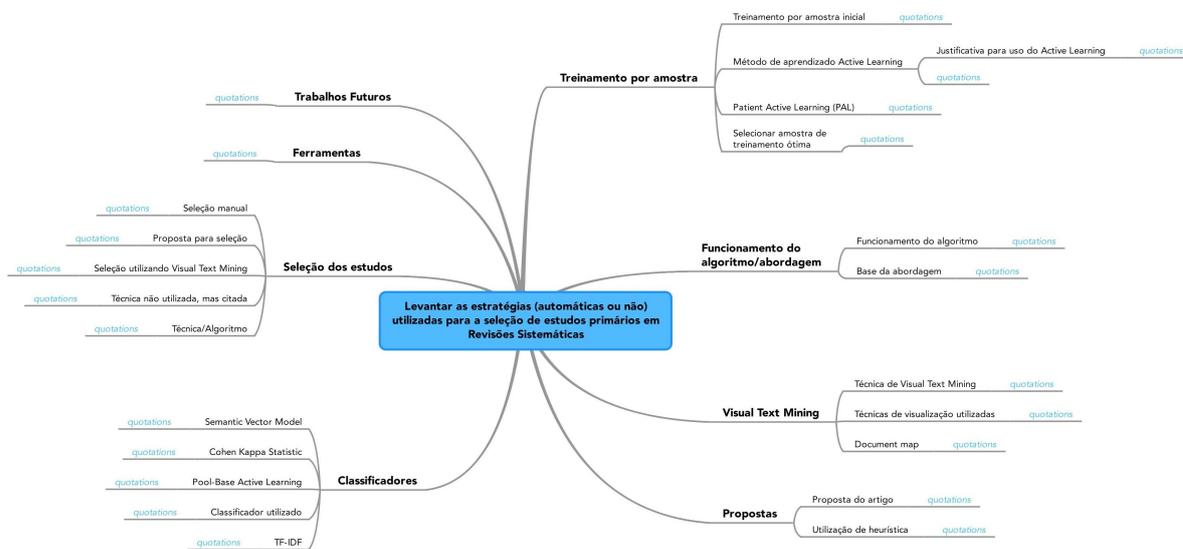


### 4.3.6 Análise e Discussão dos Resultados

De posse dos dados extraídos pelos pesquisadores, foi possível efetuar uma análise sobre os impactos gerados na extração dos dados.

Os dados extraídos por cada participante são apresentados em formato de mapa mental. O mapa mental dos participantes P1 e P3 corresponde às categorias e sub-categorias que foram adquiridas e refinadas durante a extração e sumarização dos dados. Cada categoria e sub-categoria possui associada a elas diversas *quotations*- correspondem ao dado bruto extraído e organizado na respectiva categoria. O mapa mental de P2 representa o formulário de extração refinado por ele. Cada ramo do mapa corresponde a uma pergunta do formulário que possui respostas associadas de cada estudo primário.

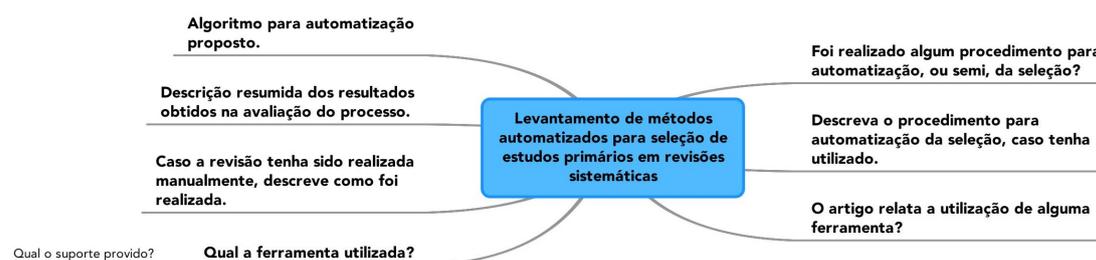
O modelo refinado dos dados extraídos por P1 pode ser visualizado na Figura 4.10. Os dados extraídos por P2 correspondem aos campos do formulário de extração, apresentados na Figura 4.11. Por fim, os dados de P3 são representados na Figura 4.12).



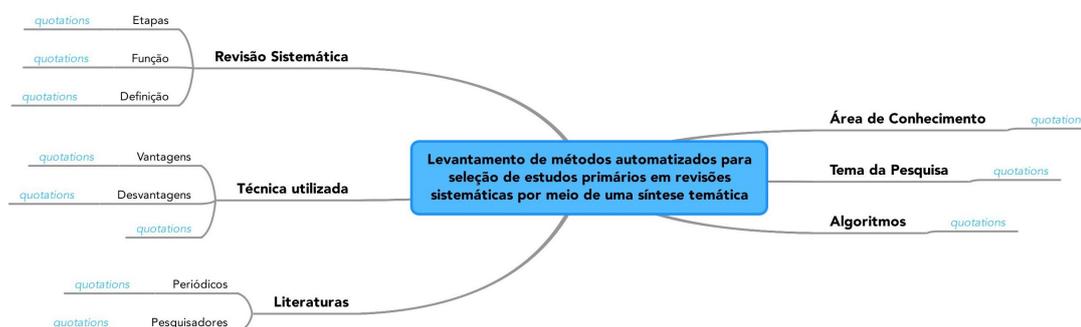
**Figura 4.10: Modelo de mapa mental obtido da extração dos dados por P1**

Como primeiro ponto a ser analisado, verificamos os dados que foram extraídos e refinados. Uma breve análise dos dados de cada pesquisador é apresentada a seguir:

- Ao analisar o mapa mental de P1 é possível ver um relacionamento de tópicos (apresentados no primeiro nível do mapa mental) e tópicos derivados de cada um deles (apresentados



**Figura 4.11: Modelo de mapa mental obtido da extração dos dados por P2**



**Figura 4.12: Modelo de mapa mental obtido da extração dos dados por P3**

no segundo nível do mapa mental), seguido pelas *quotations* relativas a cada um. Sem grande esforço por parte de quem analisa este mapa, é possível compreender o contexto da pesquisa e como alguns dos itens se relacionam uns com os outros. Em posse deste modelo, o pesquisador já tem pré preparada uma estrutura para sua sumarização e estado da arte, cabendo a ele ligar e dar ênfase aos pontos mais importantes para a pesquisa.

- Já o mapa mental de P2, correspondente a seu formulário de extração, apresenta as questões de seu formulário final. Cada questão possui uma resposta associada para cada estudo. Para realizar a sumarização dos dados o pesquisador deve correlacionar estas informações e formular o estado da arte da pesquisa.
- O formulário de P3, assim como o de P1, busca efetuar um relacionamento dos tópicos coletados. Possui uma estrutura pré-formulada para sumarização e formulação do estado da arte do tema. Quando comparado a P1, verificamos que o nível de detalhes do mapa é menor - justificável pela experiência de P3.

Avaliar a qualidade da análise dos dados é subjetivo. Porém, verificamos que a essência dos dados extraídos por P1 e P2 são similares, a principal diferença é o modo como foi conduzido e a organização de como esses dados são representados. P2 possuía domínio do tópico de pesquisa

tratado na revisão sistemática conduzida. Já P1 não dominava o tópico de pesquisa e extraiu todos os dados que julgou que poderiam vir a ser relevantes, assim como propõe a utilização da abordagem RS/AT, e, ao final, possuía um domínio maior do tópico tratado no estudo e pode refinar estes dados extraídos. Mesmo P2 tendo maior domínio, a essência dos dados extraídos de P1 e P2 foi similar. Quanto a P3, mesmo sem nenhuma experiência em revisão sistemática, na abordagem RS/AT e nem no tópico de pesquisa, chegou a uma extração dos dados considerada satisfatória. Esta análise sugere que, a utilização da abordagem RS/AT (quando conduzida por pesquisadores que tenham um conhecimento razoável dela) pode contribuir significativamente para a extração dos dados para pesquisadores que não possuem domínio do tópico que será tratado na revisão sistemática, evitando assim que dados relevantes não sejam coletados e permitindo construir o conhecimento do tópico a medida que a pesquisa é conduzida.

Outro ponto analisado foi o esforço dispendido para condução da atividade. Ressaltamos que mensuramos neste ponto esforço, não tempo. **Por esforço pode-se entender a junção da quantidade de retrabalho que o participante teria que realizar, ou seja, quanto da análise previamente realizada teria necessidade de ser feita novamente (mesmo que parte dessa análise). Visto que o foco de uma revisão sistemática é a qualidade do estado da arte obtido, o tempo dispendido na atividade não é um dado tão relevante (dentro de um limite tolerável para tal).**

No quesito esforço, a abordagem com utilização da análise temática, durante a extração de dados, dispense menor esforço do pesquisador. Apesar de, com a análise temática, o pesquisador teoricamente adiciona uma fase a sua revisão (a pré-análise dos estudos), na prática o esforço dispendido é menor.

Enquanto na abordagem tradicional é altamente recomendado a releitura dos estudos já analisados sempre que o formulário de extração de dados tem um campo adicionado ou alterado (para garantir a integridade dos dados coletados). Já com a análise temática, este esforço adicional não é dispendido visto que a coleta dos dados, quando identificados que possam ser relevantes é encorajado. Ao final do processo os dados coletados serão refinados e, se identificados como não relevantes, basta serem descartados.

**Apesar de, idealmente, o esforço dispendido com a análise temática ser menor, existe o risco tanto na abordagem tradicional quanto na com utilização da análise temática de a extração dos dados ser mal realizada e então o trabalho não atinge o grau de qualidade desejado. Porém, este fator citado é inerente de qualquer uma das abordagens utilizadas e dificilmente pode ser minimizado.**

Quanto ao modo de extração dos dados, alguns pontos valem ser ressaltados. O primeiro é o modo distinto em como é realizado - um utilizando os formulários e outro codificando os

documentos e agrupando os trechos em categorias. Ambas as abordagens necessitam de um planejamento inicial bem estruturado a fim de evitar as seguintes dificuldades que possam a vir ocorrer:

- **Omissão de evidências:** No caso da utilização do formulário de extração, os dados serão coletados de acordo com o formulário - caso este não seja adequadamente definido, ocorrerá a omissão de evidências.

Já utilizando a abordagem RS/AT, inicialmente não há uma rigorosa definição dos dados que serão coletados. São definidos objetivos, questões a serem respondidas e, até mesmo, o desenvolvimento de *seed categories* (categorias definidas inicialmente para serem utilizadas como guias na coleta). Porém, é recomendável que todo o dado que possa ser interpretado como possivelmente relevante seja coletado - ao fim da codificação, caso os dados sejam julgados como não relevantes em consenso, basta descartá-los.

- **Retrabalho:** Ao se utilizar o formulário de extração, a coleta dos dados é padronizada. Porém, caso estes não sejam corretamente definidos no início da pesquisa, dados relevantes podem ser desconsiderados ou então há a necessidade de adicionar/remover campos do formulário e revisar os estudos analisados previamente a fim de identificar se aquela evidência está contida em algum daqueles estudos.

No caso da abordagem RS/AT, o retrabalho durante a extração dos dados não ocorre - salvo exceções onde os objetivos da revisão sejam alterados, o que caracterizaria a condução de uma nova revisão.

- **Dados discrepantes:** Quando se conduz revisões colaborativas (mais de um pesquisador conduzindo a revisão), os dados coletados devem ser minimamente padronizados para manter a consistência das informações na sumarização.

Utilizando formulários de extração esta padronização é relativamente simples. Basta seguir o formulário e, havendo necessidade de alteração, efetuar uma reunião para discutir esta alteração e atualizar o formulário dos pesquisadores envolvidos. Apesar de, conceitualmente, ser uma atividade a princípio simples, esta alteração de formulário ocasionará o ponto ressaltado anteriormente - retrabalho por parte dos pesquisadores.

Já quando se trata da abordagem RS/AT, esta padronização torna-se um pouco mais complexa. Apesar de, assim como na omissão de evidências, a formulação de *seed categories* auxilie, novos dados podem ser identificados passíveis de coleta quando conduzida a revisão. Portanto, recomenda-se que efetue a coleta de dados que possam ser, de algum modo, relevantes, mesmo que estes sejam descartados na fase de sumarização.

Os mencionados pontos foram identificados neste estudo. Observa-se omissão de evidências nos dados coletados por P3. A falta de experiência com a utilização das técnicas de revisão sistemática e análise temática contribuiu negativamente neste ponto. Porém, de acordo com os objetivos estipulados, o participante resumiu satisfatoriamente os dados.

Quanto ao quesito retrabalho, P2 alterou o formulário de extração diversas vezes durante a pesquisa e, com isso, efetuou a releitura dos estudos analisados a fim de coletar evidências que pudessem ter sido omitidas.

Por fim, no quesito de dados discrepantes, notamos que os dados de P1 e P2 estão mais próximos um do outro do que os dados de P3. Nenhum dos pesquisadores se comunicaram uns com os outros durante o estudo pois não era o intuito. Observa-se que os pesquisadores mais experientes obtiveram resultados mais satisfatórios e similares quando comparados um com o outro - resultado esperado visto que pesquisadores mais experientes tendem a ter mais facilidade e habilidade na condução de um estudo do que um pouco experiente.

#### **4.3.7 Ameaças à validade**

Quanto à validade de construção, pode-se citar:

- Métricas (dificuldade de utilização, quantidade de retrabalho, consistência dos dados e utilidade da ferramenta Insight): As métricas estabelecidas podem caracterizar uma ameaça à validade, mas todas elas tiveram um propósito bem definido para serem utilizadas, conforme descrito na seleção das variáveis dependentes.

Quanto à validade interna, pode-se citar:

- Participantes: Uma ameaça é identificada é o diferente grau de conhecimento dos participantes quanto ao tema abordado na revisão deste estudo piloto. P1 tinha conhecimento do tópico, P2 era um especialista na área e P3 não possuía nenhum conhecimento. O conhecimento do tópico, ou falta dele, pode ter influenciado os resultados obtidos. Não foi possível minimizar esta ameaça.
- Treinamento: uma ameaça identificada correspondia a falta de conhecimento da abordagem RS/AT pelo participante P3. Esta ameaça foi minimizada realizando um treinamento específico para o participante.

Quanto à validade de conclusão, verifica-se:

- Análise de dados interpretativa: Parte da análise dos dados foi realizada de modo interpretativo visto que buscou compreender informações relativas ao processo da abordagem RS/AT em relação ao formulário de extração. No entanto, ressalta-se que para minimizar esta possível ameaça, todas as interpretações e hipóteses geradas foram reforçadas pela observação dos dados extraídos e entrevista com os participantes do estudo.

Quanto à validade externa, pode-se citar:

- Participantes: Os três participantes são da área de ciência da computação. O tópico de pesquisa abordado no estudo também é da área de ciência da computação. A generalização das conclusões obtidas neste estudo deve ser avaliada para outros perfis de participantes.

## 4.4 Experimento

O objetivo deste experimento foi avaliar a abordagem RS/AT e como os recursos adicionados na ferramenta Insight contribuem para a condução da extração dos dados em revisão sistemática. A definição do experimento é apresentada na Tabela 4.5.

**Tabela 4.5: Definição do Experimento I**

<b>Objeto analisado</b>	A abordagem RS/AT
<b>Com o propósito de</b>	Avaliar a utilização da análise temática e do formulário de extração de dados no contexto de revisão sistemática
<b>Com respeito a</b>	Dificuldade de utilização, quantidade de retrabalho, dados recuperados e utilidade da ferramenta Insight.
<b>Do ponto de vista dos</b>	Pesquisadores
<b>No contexto de</b>	Alunos de pós graduação

### 4.4.1 Seleção do Contexto

O experimento foi realizado com alunos de pós-graduação em Ciência da Computação da Universidade Federal de São Carlos. Os participantes eram interessados sobre técnicas de revisão sistemática.

Assim como no estudo piloto anterior, de acordo com Wohlin et al. (2000), o contexto desse experimento pode ser caracterizado de acordo com as quatro dimensões, como:

- Processo: o processo é considerado *Off-line* para este experimento pois os participantes

não foram entrevistados durante toda a condução do experimento, apenas ao final de cada etapa.

- Participantes: o contexto dos participantes eram estudantes de pós-graduação.
- Realidade: a realidade do experimento é caracterizada como modelada pois as ferramentas utilizadas não foram desenvolvidas pela indústria e o estudo foi executado em ambiente acadêmico.
- Generalidade: a generalidade dos resultados desse experimento é classificada como específica - os resultados do experimento não podem ser generalizados para outros contextos.

#### 4.4.2 Seleção das Variáveis

Para esse experimento, assim como no estudo piloto anterior, foi identificada uma variável independente, denominada Técnica de Extração de Dados, a qual representa a abordagem de extração de dados utilizada pelos participantes. A variável independente possui dois tratamentos:

- Formulário de extração: representada pela aplicação do formulário de extração de dados durante a condução da revisão sistemática, utilizando a ferramenta StArt.
- RS/AT: representada pela aplicação da abordagem RS/AT, proposta neste trabalho, para extração dos dados utilizando a ferramenta Insight.

Quanto às variáveis dependentes, foram identificadas quatro métricas mencionadas na Tabela 4.6. Essas métricas foram selecionadas com o objetivo de validar as evidências estabelecidas no estudo piloto descrito na Seção 4.3. Exceto a métrica dados recuperados, as demais foram coletadas por meio dos questionários 1 e 2 apresentados no Apêndice C. Essas métricas são explicadas a seguir:

- Dificuldade de utilização: essa métrica deve caracterizar se a abordagem RS/AT é simples de usar, de forma a se tornar um método alternativo para extração de dados em revisão sistemática. Essa medida foi estabelecida por meio de um valor numérico em uma escala de 1 (fácil de utilizar) a 5 (muito difícil de utilizar) e por meio de comentário dissertativo explicando a opinião sobre o uso da abordagem (questões 3 e 4 do Formulário 1 do Apêndice C).

- Quantidade de retrabalho: essa métrica mede quantas vezes o formulário de extração de dados foi atualizado pelos participantes, provocando releitura do estudos primários que já tinham passado pela extração de dados. Ressalta-se que com a abordagem RS/AT esse tipo de retrabalho não ocorre, visto que as informações (dados) relevantes dos estudos primários são extraídos por meio da análise temática à medida que elas surgem. Essa medida foi estabelecida por meio de um valor numérico fornecido pelo participante, que retratou o número de atualizações que ele realizou no formulário de extração definido inicialmente (questão 2 do Formulário 1 do Apêndice C).
- Utilidade da ferramenta Insight: essa métrica foi baseada no modelo TAM (*Technology Acceptance Model*), proposto por Davis, Bagozzi e Warshaw (1989), que busca identificar por quê uma tecnologia é aceita ou não pelo usuário com base em dois conceitos: percepção sobre utilidade (se o usuário acredita que a tecnologia auxilia no desempenho da atividade) e percepção sobre facilidade de uso (o quanto a tecnologia minimiza o esforço para execução da tarefa). Essa medida foi estabelecida por meio das respostas dos participantes, às questões do Formulário 2, apresentado no Apêndice C).
- Consistência dos dados: essa métrica representa quais dados foram recuperados por cada um dos participantes. Estabeleceu-se como dados relevantes o universo de dados extraídos por todos os participantes. Essa medida foi estabelecida com a aplicação da métrica *Recall* (POWERS, 2011) no resultado de cada um dos participantes.

### 4.4.3 Projeto Experimental

O projeto experimental deste estudo foi elaborado com oito participantes, denominados P1, P2, P3, P4, P5, P6, P7 e P8, separados em dois grupos, denominados G1 e G2, com quatro participantes cada. O grupo G1 utilizou a abordagem RS/AT para extração dos dados. Já o grupo G2, utilizou formulário de extração para extração dos dados. Cada participante realizou o experimento individualmente.

Os grupos foram formados de acordo com a experiência dos participantes em revisão sistemática e análise temática, com base em uma escala de 1 a 5 (sendo 1 para pouca experiência com a técnica até 5 para muito experiente). Esta experiência dos participantes foi medida com base no formulário de caracterização, disponibilizado online aos participantes (disponível no Apêndice B). A descrição dos grupos e participantes é apresentada na Tabela 4.6.

Para tentar minimizar ameaças ao experimento, os grupos foram formados com base na experiência dos participantes em revisão sistemática. Ou seja, a média de experiência em revisão

sistemática tanto de G1 quanto de G2 é similar. Quanto ao quesito de experiência em análise temática, dois dos participantes mais experientes foram selecionados para G1 pois o grau de experiência em análise temática não impactava necessariamente nos resultados do G2, os quais utilizariam a abordagem tradicional com formulário de extração de dados.

**Tabela 4.6: Caracterização dos participantes do experimento**

Grupo	Participante	Experiência em RS	Experiência em AT
G1	P1	3	1
G1	P2	4	4
G1	P3	4	4
G1	P4	3	1
G2	P5	2	1
G2	P6	4	2
G2	P7	4	2
G2	P8	3	1

#### 4.4.4 Instrumentação

Assim como no estudo piloto apresentado anteriormente, os instrumentos do experimento foram a ferramenta StArt, implementada em Java e utilizada para dar suporte à condução de revisão sistemática, e a ferramenta Insight, também implementada em Java e utilizada para dar suporte à análise qualitativa de dados.

#### 4.4.5 Preparação e Execução do Experimento

A primeira etapa do estudo experimental foi o desenvolvimento do protocolo da revisão sistemática, apresentado na Tabela 4.7, que foi definido em uma revisão sistemática já concluída por um especialista do tópico de pesquisa e adaptado para condução deste estudo específico. Esse mesmo protocolo foi utilizado tanto pelos participantes do grupo G1 quanto pelos participantes do grupo G2.

O tema da revisão foi levantar o panorama da área de teste em sistemas embarcados. Esse tema foi escolhido pois é um tema específico de uma área da Engenharia de Software, minimizando assim o viés de conhecimento prévio por parte dos participantes.

Preparado o protocolo da revisão, o experimento foi realizado em três etapas, conforme apresentado na Tabela 4.8.

A primeira etapa do experimento foi a caracterização dos participantes, realizada por meio de um formulário sobre o grau de conhecimento em revisão sistemática e análise temática. Esse

**Tabela 4.7: Dados do planejamento da revisão disponibilizado aos participantes**

<b>Objetivo</b>	Identificar abordagens de teste de software que possam ser aplicados em sistemas adaptativos. Complementarmente, caracterizar os desafios da área durante a execução das atividades de teste, assim como os tipos de falhas e dificuldades que podem ocorrer
<b>Questão principal</b>	Quais são as abordagens propostas para teste de software de sistemas adaptativos?
<b>População</b>	Literatura sobre teste de software relacionado a sistemas adaptativos
<b>Intervenção</b>	Estudos primários que propõe teste de software em sistemas adaptativos
<b>Controle</b>	Não possui
<b>Resultados</b>	Caracterização simplificada da área de teste de software em sistemas adaptativos
<b>Aplicação</b>	Desenvolvedores e pesquisadores da área

**Tabela 4.8: Atividades realizadas no experimento**

	<b>Grupo G1</b>	<b>Grupo G2</b>
<b>Etapa 1</b>	Caracterização dos participantes	
<b>Etapa 2</b>	Treinamento sobre revisão sistemática	
	Treinamento com a ferramenta Insight	Treinamento com a ferramenta StArt
	Discussão sobre dúvidas	
<b>Etapa 3</b>	Extração dos dados com a Insight (abordagem RS/AT)	Extração dos dados com a StArt (abordagem com formulário de extração)
<b>Etapa 4</b>	Discussão dos resultados	

formulário está apresentado no Apêndice A.

A segunda etapa consistiu em treinar todos os participantes em revisão sistemática, elencando objetivos e método de extração que seria utilizado (abordagem RS/AT ou formulário de extração). Ainda nessa etapa, cada participante recebeu o treinamento presencial sobre a ferramenta que utilizaria (Insight ou StArt) e ainda teve a oportunidade de tirar dúvidas sobre os assuntos tratados.

Como os participantes do experimento já conheciam a revisão sistemática, o treinamento consistiu do ensino presencial a cada um sobre o funcionamento da ferramenta que seria utilizada (StArt ou Insight) e, para os participantes do G1, treinamento presencial com o autor deste trabalho sobre a utilização da abordagem RS/AT.

A terceira etapa foi a extração dos dados realizada por cada participante individualmente. Para tal, foi disponibilizado aos participantes os seguintes materiais:

- Instalador da ferramenta: foi disponibilizada a ferramenta Insight para o grupo G1, que

realizou a extração utilizando a abordagem RS/AT, e a ferramenta StArt para o grupo G2, que utilizou a abordagem com formulário de extração.

- Protocolo da revisão sistemática (descrito na Tabela 4.7).
- Estudos selecionados para extração: foram selecionados cinco estudos primários sobre o tópico para extração de dados.

Finalmente, a quarta etapa, consistiu da discussão dos resultados na qual os participantes elencaram os pontos principais da atividade conduzida, enviaram os dados coletados para o autor deste trabalho, responderam um questionário sobre a atividade conduzida (Formulário 1 do Apêndice C) e, os participantes do grupo G1, responderam também um questionário sobre a utilização da ferramenta Insight (Formulário 2 do Apêndice C).

#### 4.4.6 Análise e Discussão dos Resultados

Com base nos dados coletados, apresentam-se as análises conduzidas para cada uma das métricas utilizadas neste estudo.

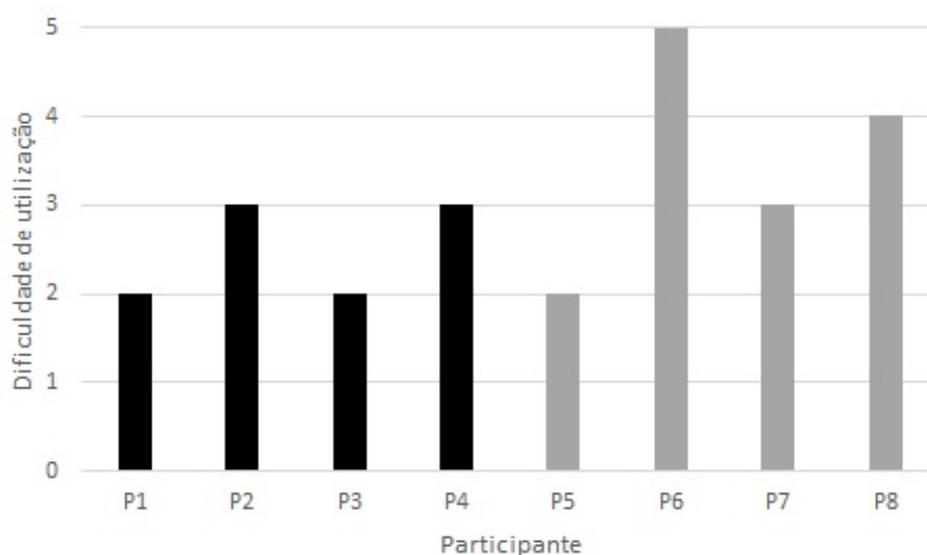
- **Dificuldade de utilização**

Ponto que deve ser analisado cuidadosamente é a dificuldade de utilização de cada abordagem encontrada pelos participantes do experimento. O grau de dificuldade foi respondido pelos participantes no formulário de avaliação do experimento (Apêndice C).

A Figura 4.13 ilustra o gráfico que relaciona cada participante (eixo X) e o grau de dificuldade encontrado por ele para condução da atividade (eixo Y). O grau de dificuldade foi definido em uma escala de 1 a 5, sendo 1 para "Não encontrou dificuldade" e 5 para "Muita dificuldade".

Os participantes do grupo G1 (P1, P2, P3 e P4) utilizaram a abordagem RS/AT. A média do grau de dificuldade encontrado pelo grupo G1 foi de 2,5 na escala de 1 a 5. Já os participantes do grupo G2 (P5, P6, P7 e P8), utilizaram o formulário de extração e a média do grau de dificuldade foi de 3,5.

Ao comparar a dificuldade média do grupo G1 com a do grupo G2, verificamos que os participantes que utilizaram a abordagem RS/AT (grupo G1) encontraram menos dificuldade na execução da atividade. Esse dado foi corroborado com observações textuais feitas pelos participantes no formulário de avaliação (Apêndice C).



**Figura 4.13: Dificuldade de utilização descrita por cada participante**

Assim, pode-se destacar do G1, que usou RS/AT, os seguintes comentários: P1 citou que a atividade foi de fácil condução, porém, encontrou um pouco de dificuldade ao definir quais seriam os códigos mais adequados de serem atribuídos às passagens que julgava relevante. P4 citou que realiza revisão sistemática com formulário de extração, porém, achou a abordagem RS/AT muito prática.

Do grupo G2, que usou formulário de extração, pode-se destacar: P6 descreveu que foi difícil chegar a uma definição sobre o formulário de extração pois, a medida que analisava os estudos, novos tópicos eram identificados como relevantes e precisavam ser coletados. Finalmente, P8 cita que a principal dificuldade encontrada foi construir um formulário de extração adequado e, ao final, alguns estudos não se adequavam ao formulário desenvolvido. Os outros participantes não relataram as dificuldades/comentários sobre a abordagem utilizada.

- **Quantidade de retrabalho**

Outro ponto analisado nesse experimento foi a quantidade de retrabalho que os participantes tiveram. Esta questão foi direcionada aos participantes do grupo G2, que desenvolveram e utilizaram formulário de extração ao invés de análise temática.

Sempre que um campo é adicionado ou alterado no formulário, é recomendado que sejam analisados novamente os estudos já analisados a fim de verificar se esse campo não foi omitido das análises realizadas anteriormente. Dentre os participantes do G2, com exceção do participante P8, todos os outros alteraram o formulário de extração da seguinte forma: P5 alterou duas vezes, P6 alterou três vezes e P7 alterou uma vez.

Verificamos que, mesmo os participantes do experimento trabalhando com um conjunto reduzido de estudos primários (total de cinco estudos para análise), na maior parte dos casos foi necessário alterar o formulário de extração. Ressalta-se que os participantes não possuíam auxílio de um especialista para construção do formulário e apenas o participante P6 possuía conhecimento técnico do tema da revisão conduzida. No entanto, mesmo P6 alterou o formulário três vezes.

Independentemente das restrições mencionadas, o retrabalho ocasionado pela alteração do formulário de extração é recorrente em revisão sistemática mesmo quando o pesquisador é especialista no tópico tratado. Esse fator pode ocorrer pois, durante a pesquisa, novos conhecimentos são adquiridos. Com a abordagem RS/AT, propõe-se que todos os dados considerados relevantes sejam coletados durante a extração de dados com a análise temática. Ao efetuar o correlacionamento das categorias, conforme proposto na análise temática, caso algum dado seja identificado como não relevante, ele é descartado para a etapa de sumarização. Ao proceder dessa forma, coletando todos os dados considerados relevantes, no momento em que um estudo primário está sendo analisado e lido, reduz-se o retrabalho, pois não há necessidade de analisar novamente estudos primários que já haviam sido lidos e analisados.

#### • Utilidade da ferramenta Insight

A utilidade da ferramenta Insight foi coletada com base no modelo TAM (DAVIS; BAGOZZI; WARSHAW, 1989), mencionado anteriormente, do ponto de vista dos participantes do grupo G1, que foi o que utilizou a ferramenta. A avaliação dos participantes foi coletada por meio do Formulário 2 do Apêndice C.

As respostas coletadas são descritas a seguir:

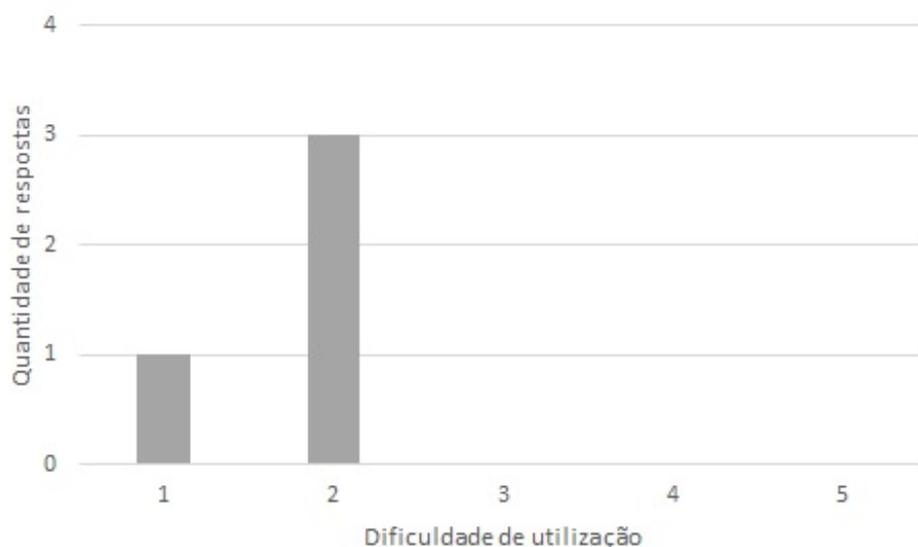
##### 1. *Foi fácil aprender a utilizar a Insight?*

Para esta questão todos os participantes que utilizaram a ferramenta assinalaram a alternativa "sim, foi fácil aprender a utilizar a ferramenta Insight".

##### 2. *Considero a Insight fácil de utilizar?*

Essa questão foi respondida por meio de uma escala numérica de 1 a 5 (sendo 1 para fácil de utilizar e 5 para difícil de utilizar).

A resposta dos participantes é ilustrada no gráfico da Figura 4.14 e, como pode ser observado, não houve resposta com valor acima de 2.



**Figura 4.14: Dificuldade de utilização da ferramenta Insight**

3. *Utilizar a Insight facilitou a execução da atividade?*

A questão 3 buscou avaliar a utilidade da ferramenta Insight para conduzir a análise temática.

Todos os participantes responderam que "Sim, a ferramenta ajudou a execução da atividade".

4. *Quão útil você considera a Insight para extração dos dados em revisão sistemática?*

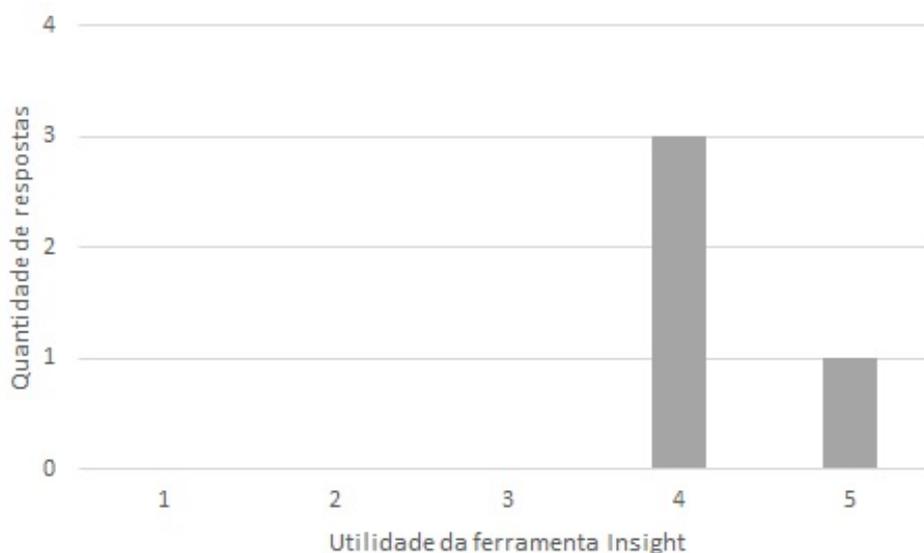
Finalmente, a questão 4 buscou avaliar o grau de utilidade da ferramenta Insight, no contexto da abordagem RS/AT, para extração dos dados em revisão sistemática.

As respostas sobre a utilidade da ferramenta foram dadas em uma escala de 1 a 5, sendo 1 para "Não considero útil" e 5 para "Muito útil". A Figura 4.15 ilustra a resposta dos participantes.

• **Dados recuperados**

A última variável dependente diz respeito aos dados extraídos dos estudos primários. Assim, analisaram-se os dados extraídos por cada participante a fim de elencar todos os dados relevantes identificados por todos os participantes. A métrica utilizada para calcular os dados recuperados foi a métrica *Recall*, descrita por Powers (2011), que é utilizada para análise de dados relativos à recuperação de informações.

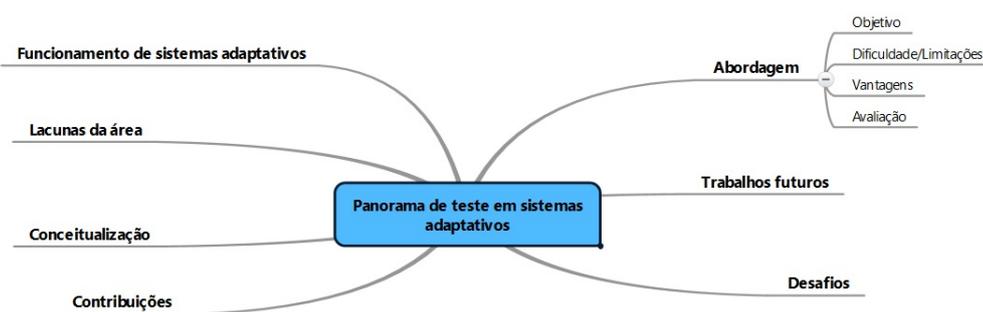
Primeiramente, apresenta-se o conjunto de dados obtido por cada participante individualmente. Para ilustrar esses dados, foi gerado um mapa mental que representa, respectivamente,



**Figura 4.15: Utilidade da ferramenta Insight para extração dos dados em revisão sistemática**

as categorias desenvolvidas pelos participantes do grupo G1, que utilizou a abordagem RS/AT, e os campos do formulário de extração confeccionado pelos participantes do grupo G2, que utilizou formulário de extração. No nó raiz é apresentado o tema da pesquisa realizada; os ramos que saem desse nó, no nível 1, representam as categorias para G1 e os campos do formulário de extração para G2. Os ramos no nível 2, que saem dos ramos do nível 1, representa as subcategorias no caso de G1 ou subcampos no caso de G2.

As Figura 4.16 a 4.19 representam, respectivamente, os participantes P1 a P4; as Figura 4.20 a 4.23 representam, respectivamente, os participantes P5 a P8, em formato de mapa mental.



**Figura 4.16: Modelo de mapa mental obtido da extração dos dados por P1**

Uma vez coletados os dados de cada participante, para possibilitar a análise e comparação dos mesmos, definiu-se um padrão de nomenclaturam apresentado na Tabela 4.9, para os dados que, semanticamente, representavam a mesma coisa. Assim, a coluna um da tabela contém o nome padrão para o tópico tratado, a coluna dois contém uma descrição do que aquele dado representa (sua semântica) e, as colunas três a dez, contêm as categorias ou campos exatamente

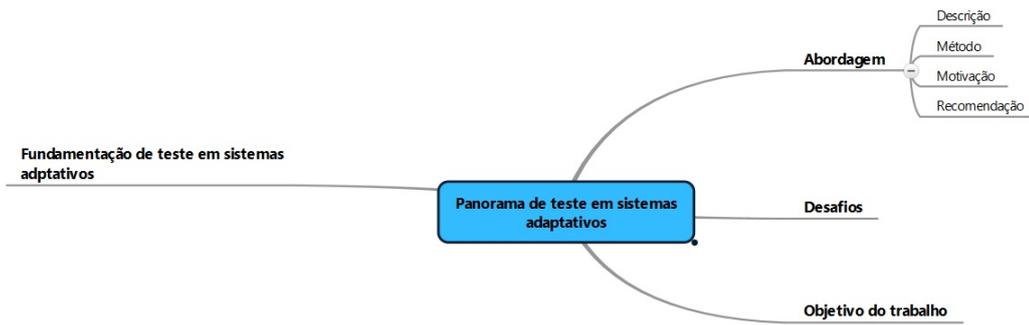


Figura 4.17: Modelo de mapa mental obtido da extração dos dados por P2

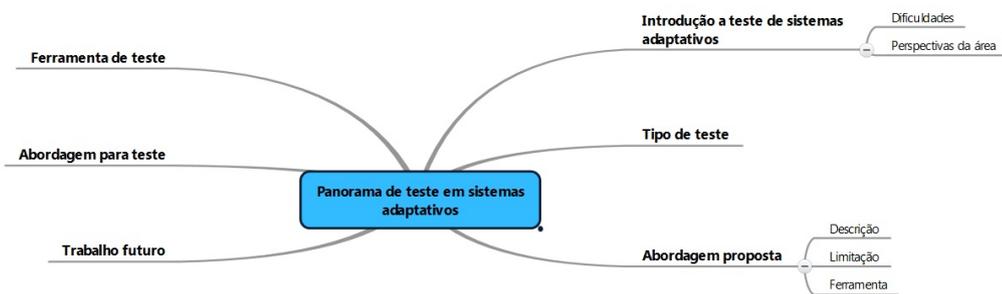


Figura 4.18: Modelo de mapa mental obtido da extração dos dados por P3

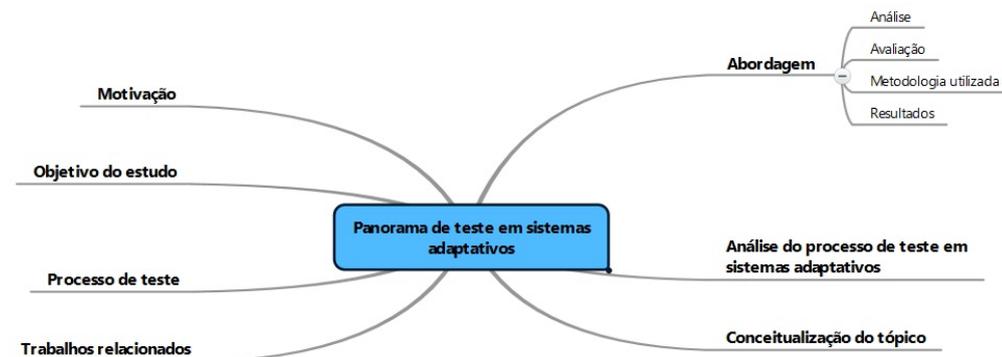


Figura 4.19: Modelo de mapa mental obtido da extração dos dados por P4

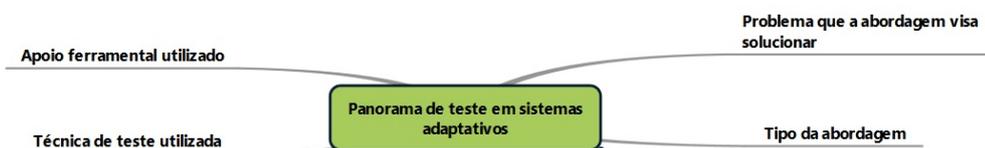


Figura 4.20: Modelo de mapa mental obtido da extração dos dados por P5

como forma extraídos pelos participantes P1 a P8, respectivamente. Nesses casos as categorias ou campos são os mais externos do mapa mental, podendo ser, portanto, do nível 1 ou do nível 2.



Figura 4.21: Modelo de mapa mental obtido da extração dos dados por P6

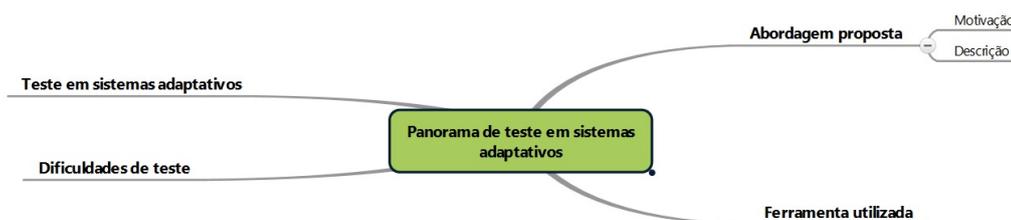


Figura 4.22: Modelo de mapa mental obtido da extração dos dados por P7

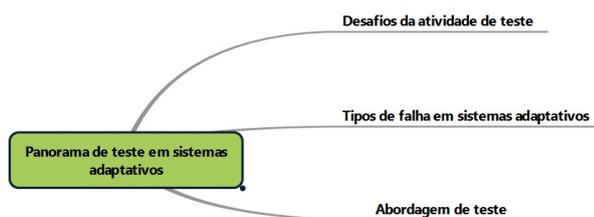


Figura 4.23: Modelo de mapa mental obtido da extração dos dados por P8

Para exemplificar essa padronização pode-se citar o seguinte exemplo, correspondente à linha três da Tabela 4.9: os participantes P1, P2, P3, P4 e P6 coletaram dados que descrevem as limitações e dificuldades relativas à utilização da abordagem proposta no estudo primário. Nesse caso, P1 atribuiu esses dados à categoria "Abordagem - Dificuldade/Limitações da abordagem", P2 atribuiu à categoria "Abordagem - Recomendações", P3 à categoria "Abordagem - Limitação", P4 à categoria "Abordagem - Avaliação" e P6 definiu um campo de seu formulário de extração intitulado "Estudo realizado - informações adicionais". Como semanticamente todas essas informações correspondem às limitações do trabalho, definiu-se a nomenclatura padrão "Limitações do trabalho", na coluna um, conforme ilustrado na Tabela 4.9. O mesmo foi feito para todos os outros dados extraídos, com o objetivo de identificar-se o total de dados diferentes que foram extraídos do conjunto dos cinco artigos tratados nessa revisão.

Após essa padronização foi possível efetuar o cálculo do *Recall* de cada participante. O *Recall* foi calculado tomando como base que todos os dados extraídos pelos participantes são dados relevantes e constituem o universo de informações, num total de 17, que deveriam ter sido extraídas por cada um. O *Recall* de cada participante é apresentado na Tabela 4.10, sendo

**Tabela 4.9: Padronização da nomenclatura dados extraídos pelos participantes**

Nome padronizado	Descrição do conteúdo	P1	P2	P3	P4	P5	P6	P7	P8
Descrição da abordagem	Dados descrevendo quais os objetivos que a abordagem e o estudo possuem	Abordagem - Objetivo	Abordagem - Descrição	Abordagem - Descrição	Objetivo do estudo	Tipo da abordagem	Estudo realizado - Resumo	Abordagem proposta - Descrição	Abordagem de teste
Limitações do trabalho	Dados descrevendo as limitações, dificuldades e recomendação de contexto para utilização da abordagem	Abordagem - Dificuldade/Limitações da abordagem	Abordagem - Recomendação	Abordagem - Limitação	Abordagem - Avaliação	X	Estudo realizado - Informações adicionais	X	X
Análise da abordagem	Dados relativos a como foi feita a avaliação da abordagem	Abordagem - Avaliação	X	X	Abordagem - Resultados Abordagem - Análise	X	X	X	X
Trabalhos futuros	Dados descrevendo trabalhos futuros no tópico de pesquisa	Trabalhos Futuros	X	X	X	X		X	X
Motivação do trabalho	Dados que descrevem as motivações de realização do trabalho	Desafios	Desafios	X	Motivação	Problema que a abordagem visa solucionar	Desafios	Abordagem proposta - Motivação	Desafios da atividade de teste
Contribuições	Dados descrevendo quais contribuições o trabalho apresenta	Contribuições	Abordagem - Recomendação	X	X	X		X	X
Conceitualização	Dados relativos a conceitualização do tema de teste em sistemas adaptativos	Conceitualização	X	X	Conceitualização do tópico	X	X	Dificuldade de teste	Tipos de falha em sistemas adaptativos
Teste em sistemas adaptativos	Dados sobre o que é teste em sistemas adaptativos	Fundamentação de teste em sistemas adaptativos	Fundamentação de teste em sistemas adaptativos	Introdução a teste de sistemas adaptativos - Perspectivas da área	Análise do processo de teste em sistemas adaptativos			Teste em sistemas adaptativos	X
Lacunas da área	Dados que apresentam lacunas da área de pesquisa	Lacunas da área	X	X	X	X	X	X	X
Funcionamento de sistemas adaptativos	Dados descrevendo como funcionam sistemas adaptativos	Funcionamento de sistemas adaptativos	X	X	X	X	X	X	X
Ferramenta de teste	Dados sobre ferramentas de teste existentes	X	X	Ferramentas de teste	X	Apoio ferramenta utilizado	X	X	X
Metodologia utilizada	Dados sobre a metodologia utilizada no trabalho	X	Abordagem - Método	Abordagem - Ferramenta	Abordagem - Metodologia utilizada	Técnica de teste utilizada	Estudo realizado - Tipo do estudo	Ferramenta utilizada	X
Trabalhos relacionados	Dados sobre trabalhos relacionados com o estudo primário analisado	X	X	X	Trabalhos relacionados	X	X	X	X
Processo de teste	Dados sobre processo de teste em geral	X	X	X	Processo de teste	X	X	X	X
Ameaças à validade do estudo	Dados sobre possíveis ameaças à validade que o estudo apresenta	X	X	X	X	X	Estudo realizado - ameaças à validade	X	X

que ele representa a quantidade de dados relevantes extraídos pelo participante em relação ao universo de informações.

Verifica-se que o participante que obteve o valor mais alto de *Recall* foi o participante P1, que utilizou a abordagem RS/AT, com valor de 0,59. Dos participantes do grupo G2, que utilizaram formulário de extração, os participantes que obtiveram o valor mais alto foram P6 e P7, ambos com 0,29. Ressalta-se que um desses participantes de G2, P6, é justamente quem tinha maior conhecimento do tema abordado na revisão. Constata-se também que, o valor mais alto obtido por dois participantes de G2 é igual ao valor mais baixo obtido por um participante de G1 (P3 que também obteve o valor de 0,29).

Tabela 4.10: *Recall* dos participantes

Participante		Total de dados coletados por todos os participantes	Quantidade de dados coletados pelo participante	<i>Recall</i>	Média do Grupo
G1 Abordagem RS/AT	P1	17	10	0,59	0,44
	P2	17	6	0,35	
	P3	17	5	0,29	
	P4	17	9	0,53	
G2 Formulário de extração	P5	17	4	0,24	0,25
	P6	17	5	0,29	
	P7	17	5	0,29	
	P8	17	3	0,18	

Com base nos dados coletados nesse experimento, verifica-se que a quantidade de dados recuperados com a abordagem RS/AT foi maior do que com a utilização do formulário de extração. A média do *Recall* do grupo G1 foi de 0,44 enquanto que a média do grupo G2 foi de 0,25.

#### 4.4.7 Ameaças à validade

As ameaças à validade deste experimento são similares às ameaças do estudo piloto descrito na Seção 4.3.

Quanto à validade de construção, pode-se citar:

- Métricas (dificuldade de utilização, quantidade de retrabalho, dados recuperados e utilidade da ferramenta Insight): As métricas estabelecidas podem caracterizar uma ameaça à validade, mas todas elas tiveram um propósito bem definido para serem utilizadas, conforme descrito na seleção das variáveis dependentes.

Quanto à validade interna, pode-se citar:

- Participantes: Quanto ao tópico da pesquisa, apenas um participante tinha conhecimento do tema. Assim, a falta de conhecimento no tema pode ter influenciado os resultados obtidos.
- Treinamento: Outra possível ameaça é o grau de conhecimento na abordagem RS/AT ou no uso do formulário de extração de dados. Nesse caso, os participantes receberam treinamento nas duas formas de extração que foram utilizadas no estudo, conforme a abordagem que iria utilizar.

Quanto à validade de conclusão, verifica-se:

- **Análise de dados interpretativa:** Como a maioria dos dados coletados no estudo foram dados qualitativos, a interpretação sobre esses dados pode caracterizar uma ameaça. No entanto, houve também alguns dados quantitativos, cuja análise sinalizou na direção das conclusões dos dados qualitativos, mostrando que a abordagem RS/AT é factível de ser utilizada.

Quanto à validade externa, pode-se citar:

- **Participantes:** Todos os participantes são da área de ciência da computação. O tópico de pesquisa abordado no estudo também é da área de ciência da computação. A generalização das conclusões obtidas neste estudo deve ser avaliada para outros perfis de participantes.

## 4.5 Considerações finais

Nesse Capítulo apresentou-se os estudos realizados que auxiliaram na avaliação e construção da abordagem RS/AT. O desenvolvimento dessa abordagem mencionada visa contribuir para a condução da revisão sistemática, fornecendo a pesquisadores mais um método para extração de dados em revisão sistemática.

Com o *survey* descrito na Seção 4.2, realizou-se o levantamento de possíveis interessados em utilizar a abordagem RS/AT. Com base nos dados apresentados pode-se verificar que existem interessados principalmente na área de computação. Visto a busca por avaliar evidências, especialmente da área de Engenharia de Software Baseada em Evidências, novos métodos para contribuir na análise dessas evidências são encorajados.

O Estudo de Caso, descrito na Seção 4.3, permitiu avaliar-se qualitativamente o processo estabelecido pela abordagem RS/AT, identificando vantagens e desvantagens da abordagem. Avaliou-se os pontos relativos à omissão de evidências, retrabalho e consistência dos dados extraídos. Verificou-se que, mesmo participantes com pouco conhecimento no tema da revisão, ou até da abordagem RS/AT, conseguem realizar uma extração de dados satisfatória utilizando a abordagem. Este fator pode-se ser justificado pelo caráter exploratório da abordagem que coleta os dados a medida que são identificados, diminuindo também o retrabalho e a omissão de evidências.

Como último ponto, no experimento descrito na seção 4.4, realizou-se com um conjunto

---

maior de participantes uma comparação do método de extração proposto na abordagem RS/AT em relação ao formulário de extração e uma avaliação da ferramenta Insight. Verificou-se que a quantidade de dados extraídos com a abordagem proposta é maior em relação ao formulário de extração. Identificou-se também que a ferramenta Insight contribui para essa extração de dados, sendo assim um suporte computacional importante para os pesquisadores utilizarem.

# Capítulo 5

## CONCLUSÃO

---

---

Este trabalho apresentou a abordagem RS/AT que utiliza análise temática em revisões sistemáticas para auxiliar a extração dos dados, o que pode impactar na qualidade da sumarização. A aplicação dessa abordagem foi viabilizada com o suporte computacional de duas ferramentas desenvolvidas em outros trabalhos deste grupo de pesquisa. Essas ferramentas são a StArt (ZAMBONI et al., 2010), que dá suporte à condução de revisões sistemáticas e a Insight (GASTALDI et al., 2015), que dá suporte à aplicação de análise qualitativa, em particular da técnica *coding*. Apesar de já implementadas as ferramentas, elas sofreram adaptações para que pudessem apoiar a abordagem RS/AT.

Uma vez definida a abordagem e viabilizado o suporte computacional, alguns estudos foram realizados com o objetivo de avaliar sua viabilidade de aplicação. Os resultados indicaram que a abordagem RS/AT pode contribuir para a extração dos dados de uma revisão sistemática, contribuindo, conseqüentemente, para a Engenharia de Software Baseada em Evidências, visto que os tópicos investigados na revisão sistemática podem ser melhor caracterizados dada a qualidade do dado extraído e da sumarização que pode ser elaborada com eles. Assim, do ponto de vista da Engenharia de Software Baseada em Evidências, pode-se dizer que técnicas, metodologias, processos, ferramentas, etc poderiam, eventualmente, ser melhor caracterizados pela e para a comunidade da área de Engenharia de Software.

Retomando as questões levantadas no Capítulo 1, ao final deste trabalho elas puderam ser respondidas da seguinte forma:

1. *Quais etapas da revisão sistemática a análise temática pode auxiliar?*

Visto que a análise temática é uma técnica para análise de dados, auxilia a revisão sistemática nas etapas de extração e, possivelmente, sumarização dos dados, uma vez que dados mais ricos podem levar a uma sumarização de maior qualidade. Assim, considera-se no

âmbito deste trabalho, que a abordagem RS/AT pode ser utilizada como uma alternativa aos formulários de extração, conforme descrito no Capítulo 3.

2. *Quando utilizada as técnicas de análise temática e revisão sistemática em conjunto (ou seja, a abordagem RS/AT) ainda se faz necessário utilizar os formulários de extração de dados?*

Com base nos resultados obtidos, principalmente nas análises apresentadas nas Seções 4.3 e 4.4, percebe-se que o uso da análise temática promove no pesquisador um entendimento mais profundo dos estudos primários e uma diversidade maior de dados relevantes que são percebidos na leitura desses estudos. Assim, considera-se que o uso da análise temática poderia substituir a elaboração do formulário de extração de dados, uma vez que, além de mais dados relevantes coletados, também há a questão do retrabalho, que não se faz necessário quando a análise temática é empregada. Ressalta-se que essas premissas devem ser exploradas com mais rigor em estudos experimentais futuros.

3. *As ferramentas StArt e Insight, desenvolvidas por outros trabalhos deste grupo de pesquisa, são passíveis de uma utilização conjunta para dar suporte à utilização da análise temática na condução de revisão sistemática?*

Conforme descrito no Capítulo 3 deste trabalho, as ferramentas StArt e Insight sofreram as alterações necessárias para permitir o uso da análise temática na condução de revisões sistemáticas. Assim, após a seleção dos estudos da revisão sistemática, é possível exportar os dados da ferramenta StArt para a ferramenta Insight, de forma que a extração de dados seja realizada com a técnica de análise temática.

4. *Quais vantagens e desvantagens podem ser observados quando do uso da abordagem RS/AT em comparação com a condução da revisão sistemática de forma tradicional, isto é, com a extração de dados baseada no formulário de extração?*

As vantagens e desvantagens obtidos foram descritas na análise dos estudos conduzidos.

Como principais vantagens identificou-se que: (i) a análise temática pode contribuir principalmente quando o pesquisador não possui total domínio do tópico investigado na revisão sistemática; (ii) evita o retrabalho que normalmente ocorre quando se usa o formulário de extração de dados que, quando alterado, exige que se repassem todos os estudos primários que já haviam sido lidos e analisados; isso não ocorre quando se usa a análise temática.

Como desvantagens pode-se citar: (i) a colta dos dados com a análise temática pode não ser totalmente padronizada em revisões colaborativas, uma vez que a extração dos dados

é realizada de forma subjetiva e exploratória. Mesmo utilizando-se um conjunto inicial de categorias, durante a aplicação da análise temática novas categorias podem surgir em decorrência do caráter exploratório mencionado. (ii) outro ponto que pode caracterizar uma desvantagem, e que deve ser explorado com maior rigor, é o fato de se extraírem muitos dados durante a aplicação da análise temática, que podem ser identificados ao final como não relevantes para caracterizar o tema explorado.

## 5.1 Contribuições e limitações

Como principais contribuições deste trabalho pode-se listar:

- Versão estável da ferramenta Insight e novas funcionalidades implementadas.

No início deste trabalho a ferramenta Insight possuía sua estrutura lógica de funcionamento implementada. Porém, algumas das funcionalidades possuíam inconsistências na implementação, resolvidas no decorrer deste trabalho. Além das funcionalidades previamente implementadas, novos recursos foram adicionados, descritos no Capítulo.

- Integração dos dados da ferramenta StArt com a Insight.

Para permitir o suporte computacional da abordagem RS/At, foram realizadas alterações nas ferramentas StArt e Insight para permitir a comunicação das mesmas por meio da troca de arquivos. Assim, os estudos que já passaram pela primeira seleção, na StArt, podem ser enviados para a Insight, de modo a serem analisados com a análise temática. Finalizada a análise dos estudos, o que pode provocar a rejeição de estudos identificados como não relevantes, essas informações podem retornar à ferramenta StArt para que os dados da revisão sejam empacotados e documentados.

- Caracterização da análise temática em conjunto com revisão sistemática.

Na busca efetuada na literatura para conduzir este trabalho, poucos trabalhos descreviam como utilizar análise qualitativa em estudos secundários. Neste trabalho procurou-se deixar claro como que os processos de um e de outro podem interagir.

- Desenvolvimento da abordagem RS/AT. Com a definição da abordagem RS/AT e do suporte computacional, pode-se conduzir uma revisão sistemática adotando-se uma forma de extração de dados que, com base nos estudos realizados, indicam que os dados extraídos dos estudos primários podem ser mais ricos e, eventualmente, impactar positivamente as evidências caracterizadas na sumarização da revisão.

Quanto às limitações pode-se citar:

- Abordagem RS/AT utiliza ferramentas específicas.

A abordagem RS/AT, proposta neste trabalho, utiliza como suporte computacional as ferramentas StArt e Insight. Salienta-se que a StArt está disponível à comunidade já há alguns anos e que a Insight será disponibilizada em breve. Apesar da abordagem RS/AT ser relativamente genérica nos seus passos, e poder ser utilizada com ferramentas similares, alguns desses passos podem necessitar adaptações, dependendo dos recursos disponíveis nas ferramentas selecionadas. Um exemplo disso, são os passos da RS/AT que fazem referência a algum tipo de visualização, como a *word cloud* utilizada na pré-análise.

- Dados devem ser migrados de uma ferramenta para outra no decorrer da pesquisa.

Conforme mencionado no Capítulo 3, a integração das ferramentas StArt e Insight foi realizada por meio da troca de arquivos. Esta escolha foi feita a fim de manter as ferramentas adequadas ao propósito original das mesmas (revisão sistemática para a StArt e análise qualitativa para a Insight).

- Estudos conduzidos.

Os estudos experimentais que foram passíveis de conduzir no contexto deste trabalho foram poucos e todos eles voltados à área de computação. Assim, outros estudos precisam ser realizados, em outras áreas de pesquisa e com perfis diferentes de participantes, para que se possam tirar conclusões mais robustas acerca da abordagem RS/AT.

- Etapa de sumarização dos dados possui recursos limitados.

A etapa de sumarização dos dados nas ferramentas ainda necessita de incremento de algumas funcionalidades para facilitar o trabalho dos pesquisadores. Entre os recursos que podem ser incrementados estão um editor de texto com mais opções de formatação e um visualizador de mapas mentais. Atualmente a ferramenta Insight exporta um arquivo de mapa mental para visualização em ferramenta externa e, na atual versão da StArt, não há recurso para ler esse arquivo e transformá-lo de fato em um mapa mental.

## 5.2 **Trabalhos futuros**

As contribuições apresentadas neste trabalho servem de motivação e base para a continuidade e aprimoramento do mesmo em trabalhos futuros. Como exemplo pode-se citar:

- Explorar com mais rigor a questão de substituir a elaboração do formulário de extração de dados pela aplicação da análise temática, como proposto na abordagem RS/AT.
- Conduzir um estudo que avalie o possível retrabalho provocado na etapa de sumarização de dados com utilização da abordagem RS/AT em relação ao formulário de extração de dados, visto que muitos dados são coletados e alguns dados podem ser descartados na sumarização do estudo.
- Adicionar novas funcionalidades na Insight para visualização de dados: atualmente, a Insight disponibiliza na própria ferramenta dois tipos de visualização dos dados para o usuário: visualização do tipo Treemap para os documentos e WordClouds para os conteúdos dos documentos. Além das visualizações disponíveis na própria ferramenta, existe a opção de exportar arquivo para geração de mapas mentais, como já mencionado anteriormente. Apesar disso, outras formas de visualização poderiam ser disponibilizadas na Insight para facilitar a aplicação da análise temática. Um exemplo disso é a *Word Tree*, que é um tipo de visualização permite analisar em que contexto as palavras de interesse são utilizadas, ou seja, são precedidas ou antecedidas por qual ou quais temas e expressões.
- Plataforma Web da ferramenta: Muitos sistemas estão migrando de plataformas locais para plataforma Web. Isto é realizado por diferentes motivos, por exemplo: centralização dos dados, facilidade de comunicação entre sistemas, entre outros.

Em vista disso, o desenvolvimento de um portal que possibilite a utilização dos recursos das ferramentas StArt para conduzir estudos secundários e Insight para análise de dados é visto como um trabalho futuro promissor. Permitiria a consulta a um acervo de dados por grande quantidade de usuários, revisões colaborativas com acompanhamento de todos os envolvidos sem necessidade de troca de arquivos, grande variedade de componentes para desenvolvimento, entre outros.

## REFERÊNCIAS

---

---

- ALDERSON, S. L. et al. How patients understand depression associated with chronic physical disease—a systematic review. *BMC Family Practice*, v. 13, n. 1, p. 41, 2012.
- BADREDDIN, O. Thematic review and analysis of grounded theory application in software engineering. *Adv. Soft. Eng.*, 2013.
- BARDIN, L. *Análise de conteúdo*. Lisboa: Edições 70, 1995.
- BASILI, V. R.; ROMBACH, H. D. The tame project: Towards improvement-oriented software environments. *Software Engineering, IEEE Transactions on*, IEEE, 1988.
- BASILI, V. R.; SELBY, R. W.; HUTCHENS, D. H. Experimentation in software engineering. *Software Engineering, IEEE Transactions on*, IEEE, 1986.
- BOYATZIS, R. E. *Transforming qualitative information: thematic analysis and code development*. Londres: Sage, 1998.
- CAMPBELL, F. et al. Behavioural interventions for weight management in pregnancy: a systematic review of quantitative and qualitative data. *BMC Public Health*, 2011.
- COLEMAN, G.; O'CONNOR, R. Using grounded theory to understand software process improvement: a study of irish software product companies. *Information and Software Technology*, 2007.
- CRUZES, D. S.; DYBA, T. Recommended steps for thematic synthesis in software engineering. In: ESEM. *2011 International Symposium on Empirical Software Engineering and Measurement*. Alberta: IEEE, 2011. ISSN 1949-3770.
- DAHAN-OLIEL, N.; SHIKAKO-THOMAS, K.; MAJNEMER, A. Quality of life and leisure participation in children with neurodevelopmental disabilities: a thematic analysis of the literature. *Quality of Life Research*, Springer, 2012.
- DAVIS, F. D.; BAGOZZI, R. P.; WARSHAW, P. R. User acceptance of computer technology: a comparison of two theoretical models. *Management science, INFORMS*, v. 35, n. 8, p. 982–1003, 1989.
- DIAS, C.; FERNANDES, D. *Pesquisa qualitativa:: características gerais e referencias*. Março 2000.
- DIXON-WOODS, M. et al. Synthesising qualitative and quantitative evidence: a review of possible methods. *Journal of Health Services Research & Policy*, 2005.

- DYBÅ, T.; DINGSØYR, T. Empirical studies of agile software development: a systematic review. *Information and Software Technology*, Newton, MA, USA, 2008.
- DYBA, T.; DINGSOYR, T.; HANSSSEN, G. Applying systematic reviews to diverse study types: An experience report. In: *Empirical Software Engineering and Measurement, 2007. ESEM 2007. First International Symposium on*. Madri: IEEE, 2007.
- FABBRI, S. P. F. et al. Externalising tacit knowledge of the systematic review process. *Software, IET*, v. 7, n. 6, p. 298–307, December 2013. ISSN 1751-8806.
- FELDMAN, R.; SANGER, J. *The text mining handbook: advanced approaches in analyzing unstructured data*. [S.l.]: Cambridge University Press, 2007.
- FELIZARDO, K. R. et al. Using visual text mining to support the study selection activity in systematic literature reviews. In: *Empirical Software Engineering and Measurement (ESEM), 2011 International Symposium on*. [S.l.: s.n.], 2011.
- GASTALDI, R. et al. Insight: uma ferramenta para análise qualitativa de dados apoiada por mineração de texto e visualização de informações. *Congresso Brasileiro de Engenharia de Software*, 2015.
- GERSHON, N.; EICK, S. G.; CARD, S. Information visualization. *interactions*, ACM, New York, NY, USA, 1998.
- HANCOCK, B.; OCKLEFORD, E.; WINDRIDGE, K. *An introduction to qualitative research*. Nottingham: Trent focus group Nottingham, 1998.
- HERNANDES, E. C. M. *Apoio a condução de análise qualitativa com técnicas de visualização e mineração de texto*. Tese (Doutorado) — Universidade Federal de São Carlos, 2014.
- HERNANDES, E. M. et al. Using visualization and text mining to improve qualitative analysis. In: *Proceedings of the 16th International Conference on Enterprise Information Systems*. [S.l.: s.n.], 2014.
- KAPLAN, B.; MAXWELL, J. A. Qualitative research methods for evaluating computer information systems. In: *Evaluating the Organizational Impact of Healthcare Information Systems*. New York: Springer, 2005.
- KITCHENHAM, B. *Procedures for performing systematic reviews*. [S.l.], 2004. v. 33.
- KITCHENHAM, B.; CHARTERS, S. *Guidelines for performing systematic literature reviews in software engineering*. [S.l.], 2007.
- KITCHENHAM, B. A.; DYBA, T.; JORGENSEN, M. Evidence-based software engineering. In: *Proceedings of the 26th International Conference on Software Engineering*. Washington: IEEE Computer Society, 2004.
- MAANEN, J. V. Reclaiming qualitative methods for organizational research: a preface. *Administrative Science Quarterly*, JSTOR, v. 1, p. 520–526, 1979.
- MAFRA, S. N.; TRAVASSOS, G. H. Estudos primários e secundários apoiando a busca por evidência em engenharia de software. 2006.

- MARSHALL, C.; BRERETON, P.; KITCHENHAM, B. Tools to support systematic reviews in software engineering: A feature analysis. In: *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*. New York, NY, USA: ACM, 2014. (EASE '14), p. 13:1–13:10. ISBN 978-1-4503-2476-2. Disponível em: <<http://doi.acm.org/10.1145/2601248.2601270>>.
- MAXWELL, J. A. *Qualitative research design:: an interactive approach*. London: Sage publications, 2012.
- MYERS, M. D. Qualitative research in information systems. *Management Information Systems Quarterly*, v. 21, p. 241–242, 1997.
- OCTAVIANO, F. et al. Semi-automatic selection of primary studies in systematic literature reviews: is it reasonable? *Empirical Software Engineering*, Springer, 2014.
- ORLIKOWSKI, W. J.; BAROUDI, J. J. Studying information technology in organizations: Research approaches and assumptions. *Information Systems Research*, v. 2, n. 1, p. 1–28, 1991.
- PATERSON, B. L.; CANAM, C. *Meta-study of qualitative health research: a practical guide to meta-analysis and meta-synthesis*. London: SAGE Publications, 2001.
- PETERSEN, K. et al. Systematic mapping studies in software engineering. In: *Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering*. Swinton, UK: [s.n.], 2008.
- PETTICREW, M.; ROBERTS, H. *Synthesizing the Evidence*. [S.l.]: Oxford Publishing, 2008.
- POPE, C.; MAYS, N. Qualitative research: reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. *Bmj*, BMJ, 1995.
- POWERS, D. M. Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. Bioinfo Publications, 2011.
- SATINK, T. et al. Patients' views on the impact of stroke on their roles and self: a thematic synthesis of qualitative studies. *Archives of Physical Medicine and Rehabilitation*, Elsevier, 2013.
- SEAMAN, C. Qualitative methods in empirical studies of software engineering. *IEEE Transactions on Software Engineering*, July 1999.
- SEAMAN, C. Qualitative methods. In: SHULL, F.; SINGER, J.; SJØBERG, D. (Ed.). *Guide to Advanced Empirical Software Engineering*. [S.l.]: Springer, 2008.
- STOL, K.-J.; RALPH, P.; FITZGERALD, B. Grounded theory in software engineering research: A critical review and guidelines. In: *Proceedings of the 38th International Conference on Software Engineering*. New York, NY, USA: ACM, 2016. (ICSE '16), p. 120–131. ISBN 978-1-4503-3900-1. Disponível em: <<http://doi.acm.org/10.1145/2884781.2884833>>.
- STRAUSS, A. L.; CORBIN, J. *Pesquisa qualitativa:: técnicas e procedimentos para o desenvolvimento de teoria fundamentada*. Porto Alegre: Artmed, 2008.
- TEIXEIRA, E. B. A análise de dados na pesquisa científica: importância e desafios em estudos organizacionais. *Desenvolvimento em Questão*, 2003.

- THOMAS, J.; HARDEN, A. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, v. 8, n. 1, p. 45, 2008.
- THOMAS, J. et al. Integrating qualitative research with trials in systematic reviews. *BMJ*, v. 328, n. 7446, p. 1010–1012, 2004.
- TONG, A. et al. The motivations and experiences of living kidney donors: a thematic synthesis. *American Journal of Kidney Diseases*, v. 60, n. 1, p. 15–26, 2012.
- TONG, A. et al. Adolescent views on transition in diabetes and nephrology. *European Journal of Pediatrics*, v. 172, n. 3, p. 293–304, 2013.
- TRAVASSOS, G. *Introdução à engenharia de software experimental*. [S.l.]: UFRJ, 2002.
- WAKERMAN, J. et al. Features of effective primary health care models in rural and remote australia: a case-study analysis. *Medical Journal of Australia*, v. 191, n. 2, p. 88–91, 2009.
- WALKER, D.; MYRICK, F. Grounded theory: an exploration of process and procedure. *Qualitative Health Research*, v. 16, n. 4, p. 547–559, 2006.
- WESEL, F. v. et al. I'll be working my way back: a qualitative synthesis on the trauma experience of children. *Psychological Trauma: theory, research, practice, and policy*, 2012.
- WOHLIN, C. et al. *Experimentation in Software Engineering: An Introduction*. [S.l.]: Kluwer Academic Publishers, 2000.
- ZAMBONI, A. et al. Start uma ferramenta computacional de apoio à revisão sistemática. *Proc.: Congresso Brasileiro de Software (CBSOFT 2010)*, Salvador, Brazil, 2010.

# Apêndice A

## QUESTIONÁRIO PARA AVALIAÇÃO DE PERFIL DE USUÁRIO EM ESTUDOS SECUNDÁRIOS E ANÁLISE QUALITATIVA DE DADOS

---

---

Questionário para pesquisa de Mestrado em Ciência da Computação - Avaliação de usuários que conduzem estudos secundários, análise qualitativa ou ambas.

### 1. Qual dos seguintes perfis se adequa ao seu?

- (a) Estudante de graduação
- (b) Estudante de pós-graduação - Mestrado
- (c) Estudante de pós-graduação - Doutorado
- (d) Graduado
- (e) Mestre
- (f) Doutor
- (g) Outro

### 2. Qual seu curso?

Ex: Ciência da Computação, Terapia Ocupacional, Educação Física, etc.

### 3. Conduz ou já conduziu estudos secundários? Se sim, quais tipos?

Selecione uma ou mais opções.

- (a) Nunca conduzi
- (b) Sim, revisão ad-hoc

(c)Sim, revisão sistemática.

(d)Outro

**4.Qual seu grau de experiência em estudos secundários?**

Como classifica o conhecimento que possui na realização de estudos secundários? 1 - Não possuo até 5 - Especialista,

(a)1

(b)2

(c)3

(d)4

(e)5

**5.Como efetua a análise dos estudos primários?**

Caso conduza estudos secundários, como extrai os dados de interesse dos estudos primários? (Ex: "Utilizo Formulários de Extração", "Efetuo um resumo de cada estudo", etc)

**6.Considera relevante o desenvolvimento de novas técnicas para auxiliar a condução de estudos secundários?**

(a)Não

(b)Sim

**7.Conduz ou já conduziu análise qualitativa de dados? Se sim, utilizou alguma técnica?**

(a)Nunca conduzi

(b)Sim, sumário narrativo

(c)Sim, análise temática

(d)Sim, meta etnografia

(e)Outro

**8.Qual seu grau de experiência em análise qualitativa de dados?**

(a)1

(b)2

(c)3

(d)4

(e)5

**9. Considera relevante a aplicação de análise qualitativa em estudos secundários?**

Ao efetuar a extração dos dados de estudos primários, consideraria utilizar a abordagem qualitativa?

(a) Não considero relevante

(b) Pouco relevante

(c) Relevante

(d) Muito relevante

# Apêndice B

## FORMULÁRIOS DE CARACTERIZAÇÃO PARA O EXPERIMENTO

---

---

**1. Já realizou alguma revisão sistemática?**

- (a) Sim
- (b) Não

**2. Como avalia seu grau de experiência em revisão sistemática?**

- (a) 1 - Pouca/Nenhuma
- (b) 2
- (c) 3
- (d) 4
- (e) 5 - Especialista

**3. Se já realizou alguma revisão sistemática, como realiza a extração e sumarização dos dados?**

**4. Já utilizou análise temática?**

- (a) Sim
- (b) Não

**5. Como avalia seu grau de experiência em análise temática?**

- (a) 1 - Pouca/Nenhuma
- (b) 2

(c)3

(d)4

(e)5 - Especialista

# Apêndice C

## FORMULÁRIOS DE AVALIAÇÃO DO EXPERIMENTO

---

---

### **C.1 Formulário 1 - *Feedback* de realização do experimento**

1.Nome.

2.Qual abordagem utilizou para extração de dados?

(a)Formulário de extração

(b)RS/AT

3.Caso tenha utilizado a abordagem com formulário de extração, foi necessário alterá-lo durante o estudo (adicionar/alterar/remover questões ou alternativas)? Se sim, quantas vezes aproximadamente?

4.Qual o grau de dificuldade que encontrou para conduzir a atividade?

(a)1 - Nenhuma

(b)2

(c)3

(d)4

(e)5 - Muita dificuldade

5.Relate suas considerações e possíveis comparações entre as abordagens (caso possua).

a) Caso tenha utilizado a abordagem RS/AT e já conheça a abordagem utilizando formulários de extração, trace um breve comparativo entre vantagens e desvantagens que julga possuir cada uma das abordagens. b) Caso tenha utilizado a abordagem RS/AT e não conheça a abordagem utilizando formulários de extração, descreva suas impressões e considerações sobre a abordagem. c) Caso tenha utilizado a abordagem com formulário de extração, descreva suas impressões e considerações.

## C.2 Formulário 2 - Avaliação da ferramenta Insight

Avaliação da ferramenta Insight quando utilizada no contexto da abordagem RS/AT.

### 1. Foi fácil aprender a utilizar a Insight?

Relate se o aprendizado para aprender sobre os recursos da ferramenta foi simples ou se encontrou dificuldades.

(a) Sim

(b) Não

### 2. Considera a Insight fácil de utilizar?

Uma vez que aprendeu a utilizar a ferramenta Insight, relate se utilizar a ferramenta é uma tarefa simples ou se é difícil de utilizar seus recursos.

(a) 1 - Fácil

(b) 2

(c) 3

(d) 4

(e) 5 - Difícil

### 3. Utilizar a Insight facilitou a execução da atividade?

(a) Sim

(b) Não

### 4. Quão útil você considera a Insight para extração dos dados em revisão sistemática?

(a) 1 - Não considero útil

(b)2

(c)3

(d)4

(e)5 - Muito útil