**Apêndice A – revisão sistemática da literatura**

Finalmente, os efeitos do impacto da tecnologia em termos de informações relevantes para profissionais na prática, como custo, qualidade e disponibilidade para o mercado são descritos no Resultado, do inglês **O**utcome. Compondo assim a estrutura do formato PICO para construção de String de Busca [Pai et al.].

De forma simplificada a string de Busca será composta de acordo com o diagrama de Venn disposto no Gráfico 2.1.

Gráfico 2.1 – String de Busca : = P AND I AND C AND O

As seções subsequentes especificam dados para a construção da string de busca utilizada durante a construção desta pesquisa.

População (P)

**P :=**  (“Algorithm” OR “Binary files <or> Diagrams <or> Directory <or> Development <or> Program development <or> Program project <or> Software development <or> Software Models <or> Software project <or> Software technique <or> Source code <or> System development <or> System project <or> UML <or> UML diagrams <or> XML <or> Xml documents)

Intervenção (I)

**I :=** ( Change Control <or> Comparison <or> Comparison tool <or> Conflict <or> Conflict resolution <or> Delta <or> Diff <or> Diff tool <or> Diff2 <or> Diff3 <or> Differencing tool <or> Merge <or> Merge conflict <or> Merging tool <or> Model Differences <or> Model merge <or> Model versioning <or> Software change <or> Software clone <or> Software diagram <or> Software Merging <or> Three-way merge <or> Version changes <or> Version tool

Comparação (C)

**C** := Ø

Resultado (O)

**O :=** (Changes detection <or> Clone detection <or> Conflict detection <or> Diff detection)

* + - * 1. String de Busca

Após a definição das strings referentes ao formato PICO, o protocolo de revisão a ser utilizado nesta pesquisa foi definido, de forma que **String de Busca : = P AND I AND C AND O**. Ou seja:

**String de Busca =** **(**Algorithm” OR “Binary files” OR “Diagrams” OR “Directory” OR “Development” OR “Program development” OR “Program project” OR “Software development” OR “Software Models” OR “Software project” OR “Software technique” OR “Source code” OR “System development” OR “System project” OR “UML diagrams” OR “XML documents**) AND (**Change Control” OR “Comparison” OR “Comparison tool” OR “Conflict” OR “Conflict resolution” OR “Delta” OR “Diff” OR “Diff tool” OR “Diff2” OR “Diff3” OR “Differencing tool <or> Merge” OR “Merge conflict” OR “Merging tool” OR “Model Differences” OR “Model merge <or> Model versioning” OR “Software change” OR “Software clone” OR “Software diagram” OR “Software Merging” OR “Three-way merge “ OR “Version changes” OR “Version tool**) AND (**Change detection” OR “Clone detection” OR “Conflict detection” OR “Diff detection**)**

* + - * 3. Identificação da Pesquisa

Considerações no contexto de máquinas de busca, influenciaram diretamente a construção da string proposta. Desta forma, o protocolo foi adaptado ao formato utilizado na sua máquina de busca principal, SCOPUS [SCOPUS].

Além do refinamento dos dados da string, para evitar redundâncias, foram inseridas tags específicas para execução de protocolos nesta ferramenta, visando à obtenção de resultados com maior relevância. Com isso, um novo protocolo foi elaborado conforme apresentado a seguir:

**TITLE-ABS-KEY**(("Algorithm" OR "Binary File" OR "Diagrams" OR "Directory" OR  "Development" OR "Program development" OR "Program project" OR "Software"  OR "Software development" OR "Software Models" OR "Software project" OR  "Software technique" OR "Source code" OR "System development" OR "System project" OR "UML" OR "XML") **AND** ("Diff" OR"Diff2" OR "Diff3" OR  "Differencing tool" OR "diff/merge" OR "Merge tool" OR "software merging" OR  "Merging tool" OR "Three-way merge" OR "Comparison tool" OR "Change Control" OR"Version tool" OR "Versioning tool" OR "Version changes"))  **AND**  (**LIMIT-TO**(SUBJAREA, "ENGI") OR LIMIT-**TO**(SUBJAREA, "COMP") OR  **LIMIT-TO**(SUBJAREA, "MULT"))

O protocolo descrito anteriormente foi utilizado durante a realização da Revisão Sistemática disposta neste relatório.

* + 1. Avaliação do Planejamento

O protocolo da revisão deve ser avaliado para garantir que o planejamento é viável. A partir disso é necessário testar a execução do protocolo, em diversas máquinas de busca. [Travassos e Biolchini] [Schots]

Neste trabalho a definição de artigos de controle, influenciou o processo de avaliação do planejamento. A expressão de busca foi definida através de testes de retorno dos artigos de controle, ou seja, caso a expressão de busca retornasse todos os artigos de controle, esta seria considerada ideal para realização da revisão, caso contrário a expressão seria escolhida de modo a maximizar o resultado obtido.

Além disso, a avaliação do protocolo foi realizada durante reunioes semanais sob orientação do professor Leonardo Murta. Nestas reuniões foram apresentados resultados participais e retorno do orientado foi obtido.

* 1. Execução

Todo o processo de execução, realizado em outubro de 2010, será descrito de forma minuciosa nas seções consecutivas.

* + 1. Artigos de Controle

Diversos testes foram conduzidos de forma a garantir que a expressão de busca escolhida estivesse de acordo com objetivos deste estudo.

Desta forma, foram definidos artigos relevantes para o contexto do trabalho e que serviram como artigos de controle da expressão de busca. O principal objetivo dos artigos de controle, é a validação do resultado obtido através da aplicação do protocolo na fonte de busca selecionada.

Neste trabalho, foram estabelecidos quatro artigos de controle, conforme descrito na Tabela 2.1. Após a definição das fontes a serem utilizadas, foi verificado que o artigo indexado pelo número 1 não constava nos resultados obtidos durante as etapas de avaliação. Por este motivo, um novo critério de aceite foi incluído relacionado à análise das referências bibliográficas apresentadas nos artigos selecionados. A seção 2.3.6 apresenta detalhes sobre esses critérios.

Os artigos de controle estão classificados por ano de publicação, não levando em consideração o grau de relevância dos mesmos.

|  |  |  |  |
| --- | --- | --- | --- |
| ID | Título | Autor (es) | Ano |
| 1 | An Algorithm for Differential File Comparison | J. W. Hunt  M. D. McIlroy | 1976 |
| 2 | X-diff an effective change detection algorithm for XML documents | Yuan Wang  David J.DeWitt  Jin-Yi Cai | 2003 |
| 3 | A State-of-the-Art Survey on Software Merging | Tom Mens | 2002 |
| 4 | A three-way merge for XML documents | Tancred Lindholm | 2004 |

Tabela 2.1 – Artigos de controle

* + 1. Critérios de Seleção das Fontes

O critério utilizado para seleção das fontes foi à centralização, no que diz respeito às fontes de pesquisa, devido ao curto tempo para realização da análise. Além do número de títulos armazenados na ferramenta.

* + 1. Listagem das fontes

A ferramenta SCOPUS, foi à principal fonte utilizada durante a execução deste estudo. A escolha desta máquina de busca, está fundamentada no fato deste ser o maior banco de dados de resumos e citações de literatura científica e de fontes de qualidade na web, abrangendo cerca de 18.000 títulos de mais de 5.000 editores. [SCOPUS]

O acesso ao SCOPUS foi permitido através da utilização do Proxy[[1]](#footnote-1) disponibilizado pela Universidade, com acesso restrito a alunos e funcionários da UFF. Para auxílio à busca dos documentos completos, o Google Scholar[[2]](#footnote-2) foi selecionado como máquina alternativa.

* + 1. Tipo dos Artigos

Os artigos selecionados, após a execução do protocolo, em sua totalidade, eram acadêmicos e possuíam um número de citações considerável, ampliando assim o rigor científico na construção de resultados.

* + 1. Idioma dos Artigos

Para a realização desta pesquisa, o inglês foi adotado. A escolha foi realizada, devido à compatibilidade da linguagem com conferências e periódicos internacionais, além da utilização do idioma nas fontes utilizadas para o estudo.

* + 1. Critérios de Inclusão e Exclusão dos Artigos

Os critérios de inclusão e exclusão dos artigos foram construídos de acordo com cada etapa de avaliação. Inicialmente os itens 1,2 e 3, descritos na Tabela 2.2, foram os principais critérios de exclusão adotados para a primeira etapa de avaliação. Já os itens 4 e 5 referem-se à segunda etapa, onde o documento completo é avaliado e possíveis incompatibilidades entre os temas ou inviabilidade na replicação do estudo são encontrados, ocasionando assim a exclusão do artigo.

Durante a busca do texto completo dos artigos, foram verificados que alguns deles não estavam disponíveis, de forma gratuita, para leitura. Diante desta situação emails, para os autores, foram enviados no intuito de obter estes documentos. Porém devido ao tempo reduzido, entre a solicitação e construção deste relatório, a maior parte destas solicitações não foram atendidas. Desta forma, a complementação da avaliação será realizada à medida que as respostas forem obtidas.

Em relação à inclusão de artigos, um único critério complementar foi estabelecido. Ou seja, uma das formas de busca está relacionada à análise das listas de referências bibliográficas dos artigos obtidos com a execução do protocolo. Este critério só será considerado nos artigos selecionados após a segunda etapa de revisão e adição do artigo deverá levar em consideração o grau de relevância do assunto abordado.

Conforme pode ser verificado no texto descrito anteriormente, a classificação está fundamentada na exclusão de publicações cujo tema não seja compatível com o estudo realizado. Esta estratégia foi utilizada para que a extração de características fosse maximizada.

|  |  |
| --- | --- |
| Item | Descrição |
| 1 | Incompatibilidade ou irrelevância de informações na análise do Resumo |
| 2 | Incompatibilidade ou irrelevância de informações na análise da Introdução |
| 3 | Incompatibilidade da Área de Conhecimento |
| 4 | Incompatibilidade ou irrelevância dos Temas Abordados |
| 5 | Impossibilidade de reprodução do estudo realizado |

Tabela 2.2 – Critérios de exclusão de artigos

* + 1. Processo de Seleção dos Estudos

O processo de seleção de estudos foi dividido em etapas, de forma que cada uma delas considera de maneira distinta, os critérios definidos para esta revisão. Todas as tabelas construídas com os principais dados dos artigos selecionados foram armazenadas no GoogleDocs[[3]](#footnote-3)

**1ª Etapa:** seleção dos estudos coletados: enumeração preliminar das publicações, realizada a partir da aplicação da expressão de busca na fonte principal escolhida.

**2ª Etapa:** seleção dos estudos relevantes (1º filtro): após a identificação das publicações, os resumos (*abstracts*) e as introduções foram lidas e analisadas.

**3ª Etapa:** seleção dos estudos relevantes (2º filtro): apesar de limitar o universo de busca, o filtro empregado não garante que o material coletado seja útil no contexto da pesquisa. Por isso, as publicações selecionadas na 2ª etapa devem ser lidas por completo, verificando se, de fato, atendem aos critérios definidos. [Schots]

* + - 1. Primeira Etapa de Avaliação

Após a execução do protocolo, na fonte principal de pesquisa, foram retornados 513 artigos no total. A leitura e a avaliação foram realizadas em conjunto com o aluno Eraldo Borel, do curso de graduação da Universidade Federal Fluminense, que irá construir um trabalho no mesmo tema de pesquisa.

Foram avaliados através da leitura do resumo e da introdução, os assuntos abordados, os critérios de aceite estabelecidos, principais características e possíveis críticas referentes ao estudo analisado.

O ANEXO I apresenta a lista com informações obtidas através da revisão sistemática na primeira etapa da avaliação. Já o gráfico 2 demonstra o percentual de artigos eliminados nesta fase da avaliação.

Gráfico 2.2 – Artigos Selecionados

* + - 1. Segunda Etapa de Avaliação

A partir da seleção realizada, na primeira etapa, é disposta uma nova maneira de avaliar os artigos retornados pela pesquisa na fonte. Desta forma, os documentos completos devem ser lidos, considerando critérios mais rigorosos e de maior grau de relevância. Nesta etapa, 84 artigos foram selecionados para etapa 3, e todos foram classificados conforme o Gráfico 2.3. Toda a listagem dos artigos selecionados e as características levantadas estão evidencias no ANEXO II deste documento.

Gráfico 2.3 – Classificação de Artigos

* + - 3. Terceira Etapa de Avaliação

Após leitura e análise realizada na segunda etapa de avaliação, cada publicação aprovada é examinada individualmente através da leitura completa de seu conteúdo. Nesta etapa é esperado que os dados da publicação sejam armazenados, de forma sintética, em tabelas previamente construídas. A principal finalidade é expor de maneira sucinta, dados e características relevantes da publicação, facilitando assim a posterior utilização das informações obtidas durante o processo de revisão.

Alguns elementos básicos devem ser informados para uma melhor identificação da leitura realizada, como por exemplo, título, autor (es), data da publicação, referência completa e resumo da publicação. Além disso, de forma contínua, 4 perguntas foram descritas para obtenção da sumarização dos resultados.

1. Quais métodos são utilizados para detecção de diff ?
2. Quais dados dos problemas são utilizados para a detecção de diff ?
3. Quais as características dos artefatos utilizados ?
4. Quais são as principais características consideradas durante o processo ?
   * 1. Estratégia de Extração das Informações e Sumarização dos Resultados

A partir das informações extraídas das publicações selecionadas para o estudo, a estratégia adotada era referente à extração de dados relevantes, compatíveis e passíveis de reprodução. Porém, para não restringir as informações relevantes foi decidido que todas as características com grau de relevância considerável deveriam ser expostas nas tabelas definidas na seção anterior.

O ANEXO III apresenta a avaliação individual dos arquivos selecionados e disponíveis para leitura de forma gratuita. Com a análise desta tabela foi possível responder as questões de pesquisa formuladas na seção 2.2.1.1.

* 1. Análise de Resultados

O gráfico 2.4 apresenta a divisão de tarefas durante as etapas de avaliação. No que diz respeito à construção e sumarização dos resultados, as tarefas foram divididas de forma equilibrada, tal que 41 artigos ficaram sob responsabilidade da aluna de mestrado Fernanda Floriano e a outra metade como encargo do aluno de graduação Eraldo Borel.

Gráfico 2.4 – Análise de Responsabilidades

Os resultados obtidos durante o processo de revisão foi bastante satisfatório, uma vez que foi possível listar características citadas nas publicações avaliadas. O anexo IV apresenta uma seleção de características encontradas, assim como a identificação de compatibilidades existentes. É de extrema importância citar, que o processo de revisão não foi concluído devido à ausência de artigos completos disponíveis para leitura

1. http://proxycapes.nti.uff.br/ [↑](#footnote-ref-1)
2. http://scholar.google.com.br/ [↑](#footnote-ref-2)
3. https://docs.google.com/ [↑](#footnote-ref-3)