

## **Detecção de Indícios de Plágio nos Trabalhos Acadêmicos – um Software para Auxiliar e Automatizar o Processo**

The detectation of plagiarism evidences in academic papers - A software to support and automate the process

Roseclea Duarte Medina<sup>1</sup>, Catiane Priscila Barbosa Arenhardt<sup>2</sup>

1,2 Universidade Federal de Santa Maria - UFSM - Centro de Tecnologia – Programa de Pós-graduação em Informática - Mestrado em Computação - roseclea.medina@gmail.com

### **Resumo**

O fácil acesso as informações devido a difusão da internet, possibilita o enriquecimento intelectual, mas por vezes acaba por ser um meio de usufruir do conhecimento de outrem sem mencionar seus créditos/direitos autorais, o que por fim acaba na configuração do plágio. O plágio no meio acadêmico é uma tarefa difícil de ser controlada, devido o grande número de trabalhos que são feitos por uma vasta quantidade de alunos e também pelo excesso de tarefas dos professores e pelo pouco tempo que conseguem dedicar para o controle da qualidade e autenticidade dos trabalhos. Com isso, é importante contar com softwares que auxiliem no processo de verificação de indícios de plágio, desta maneira o presente trabalho vislumbra aprimorar o método DIP – Detector de Indícios de Plágio para auxiliar o docente na verificação da autenticidade dos trabalhos.

**Palavras Chave:** softwares, indícios de plágio,

### **Abstract**

The easy access to information due to the Internet's diffusion enables the intellectual enrichment. However, oftentimes, it ends up being a way to use someone else's knowledge without mentioning them in the credits or their authorial rights. Such process is called plagiarism. Plagiarism in the academic world is a difficult task to be controlled because of the great amount of papers, which are done by a vast quantity of students, and the excess of professors' tasks e their limited time to work on the control of quality and authenticity of the papers. Thus, it is important to rely on software that supports the process of fiding plagiarism evidences. That being said, the present papers aims to improve the DIP method – Plagiarism Evidences Detector – to support the professor in the verification of the authenticity of papers

**Keywords:** softwares, Plagiarism Evidences,

## I. INTRODUÇÃO

A difusão da internet e junto a ela a quantidade de informações disponíveis formando um grande acervo virtual facilita acesso a uma infinidade de materiais, proporcionando ao usuário a possibilidade de usufruir de informações de maneira incorreta ou mal intencionada, sem dar os devidos créditos aos autores [MORAES, 2012].

No meio acadêmico, os alunos são colocados a frente da produção de materiais de pesquisa, tanto de artigos quanto de trabalhos de conclusão de curso, e os discentes por muitas vezes sentem-se inseguros com a escrita, são inexperientes, desconhecem que caracteriza cópia ilegal ou plágio e utilizam-se das informações de maneira incorreta ou ainda copiam ou compram trabalhos intencionalmente. O autor Barnbaum [2002] escreve que a falta de conhecimento do que constitui o plágio leva muitos alunos a cometê-lo inconscientemente. Se não sabe exatamente o que o plágio é, não pode evitar fazê-lo.

Dentro dessas perspectivas, é visível que aumente consideravelmente a atenção do professor em relação à autenticidade dos trabalhos, culminando no acréscimo de sua demanda de trabalho devido a maior atenção dedicada para a correção.

Comumente, os professores têm grandes quantidades de alunos e muitos acabam por não dispor de tempo para um acompanhamento mais profundo na produção acadêmica de seus alunos, assim o uso de ferramentas para auxiliar na verificação de indícios de plágio é uma alternativa para este fim.

Como já citado anteriormente, o uso de ferramentas é uma das alternativas que se pode ter com a finalidade de otimizar o tempo do professor no controle da autenticidade das informações escritas.

Com base nesses estudos apresentados, buscou-se o desenvolvimento de uma ferramenta para auxílio na verificação de indícios de plágio textual. Para tanto, partiu-se da identificação do estado da arte nessa área, bem como uma lista de ferramentas pesquisadas bibliograficamente e em seguida testes nessas ferramentas para identificação de suas funcionalidades.

O levantamento bibliográfico das ferramentas foram feitos em trabalhos já desenvolvidos em [SIBI, 2011], [LIMA e RESENDE, 2012], [PERTILE, 2011] e [SANTOS e FRANCO, 2010], e nos sites dos fabricantes [DOC COP, 2012], [EPHORUS, 2012], [ETBLAST, 2012], [FAREJADOR, 2012], [PLAGIARISMA, 2012], [PLAGIARISM.ORG, 2012], [PLAGIUM, 2012], [PLAGIUS,

2012], [PLAGIO.NET, 2012], [SCHOLARONE, 2012], [TURNITIN, 2012], [URKUND, 2012], [VIPER, 2012]. Em seguida, foram realizados testes nas ferramentas que apresentaram licença gratuita. Com base nos levantamentos bibliográficos e testes, identificou-se que as ferramentas que são de licença livre necessitam de cadastro de usuário, fazem buscas por termos similares somente na Internet, não fazem análise cruzada de arquivos inteiros, possuem verificação de arquivos com extensões limitadas e em algumas possuem limitadores de quantidade de termos permitidos para uma determinada análise ou tamanho de arquivo limitado a no máximo 300Kb.

A partir desse levantamento, foram identificadas e elencadas funções que precisam ser aprimoradas ou mesmo desenvolvidas para que uma ferramenta utilizada na detecção de indícios de plágio consiga tratar um número maior de especificidades de textos. Sendo assim, propôs-se o desenvolvimento de uma ferramenta de detecção de indícios de plágio textual em arquivos com extensões .doc, .docx, .pdf e .HTML, utilizando técnicas de stemming (extração do radical das palavras e armazenamento em uma lista) que possibilita a comparação de palavras sinônimas. Inclusão de análise de referências cruzadas, a partir do download dos documentos suspeitos que serão armazenados em diretório e este por fim formará um repositório de documentos suspeitos. Por último será realizada a comparação entre os arquivos culminando na geração do relatório apresentando os indícios de plágio.

## 2. PLÁGIO NO MEIO ACADÊMICO E FERRAMENTAS

O autor afirma em [HANDBOOK, 2009] que o plágio é definido como a utilização das palavras ou ideias de outra pessoa como se fosse seu próprio trabalho, cita-se como exemplos de plágio: copiar, traduzir um texto de um idioma para outro, parafrasear ou referenciar incorretamente. Ainda em [LACKES et. al, 2009 apud MEGEHEE and APAKE, 2008] encontra-se a afirmação que em torno de 70% das obras não são referenciadas corretamente, o que por sua vez, também, configura o ato ilegal do plágio. Segundo [SILVA, 2008], o problema do plágio dentro do meio escolar, vem desde o ensino fundamental, onde se copia textos de outrem parcialmente ou totalmente sem referenciar a fonte.

Alguns exemplos que podem ser considerados como plágios são citados a seguir [HANDBOOK, 2009; OLIVEIRA e OLIVEIRA, 2008]:

- Citação: trata-se da cópia idêntica das palavras e ideias do autor, sem fazer referência ao autor e a obra.

- Paráfrase: o escritor do trabalho transcreve com suas palavras as ideias do autor que deveria ser referenciado, com a finalidade de torná-las um pouco distintas do original, e por sua vez essas palavras não são referenciadas, caracterizando plágio.

- Resumo: é uma paráfrase mais curta, porém não segue somente as ideias do autor, o escritor também expõe as suas, contudo não referencia o autor do qual utilizou para fundamentar o texto.

- Referência: este tipo de plágio ocorre quando não se referencia a obra original e sim paráfrases presentes em uma obra secundária utilizada para a formulação ou fundamentação de um texto. Por exemplo, para construção do texto A utilizou-se as obras B e C, as quais eram resumos da obra D, a obra D deveria ter sido referenciada acompanhada das obras B e C. Este tipo de plágio ocorre quando se referencia resumos, paráfrase ou citações ao invés de referenciar a obra original.

Os tipos de plágio são definidos por [KIRKPATRICK, 2007 apud OLIVEIRA, 2007] os quais são subdivididos em:

- Plágio Direto: Cópia de uma fonte por completa sem usar citações ou referenciar o autor

- Referência Vaga ou Incorreta: Como o próprio nome traz, esse tipo de plágio acontece quando uma referência é feita de maneira incorreta, ou seja, o escritor não informa o início e o fim da referência retirada da bibliografia.

- Plágio Mosaico: este tipo de plágio é um misto de paráfrases com citações, ou seja, o escritor muda algumas palavras do autor e reformula os parágrafos, porém não faz referência a fonte, o que caracteriza o plágio.

- Plágio Extra Corporal: cópia de textos fontes externas que não a sua ou de um mesmo sujeito que é também autor de um mesmo trabalho do plagiador.

O trabalho em desenvolvimento faz análise de dois tipos de plágio, Plágio Mosaico e Extra Corporal.

## 2.1 Ferramentas para análise de indícios de plágio

Segundo [MORAES, 2004], a Internet potencializa a incidência do plágio, porém o responsável pelo ato do plágio é sem dúvida o ser humano, a Internet é apenas o instrumento de pesquisa, assim como outros instrumentos que estão disponíveis para uso. Portanto, para auxílio na detecção de indícios de plágio é importante contar com ferramentas computacionais, como as que serão apresentadas na Tabela 1, abaixo, as quais também foram estudadas em [LIMA e RESENDE, 2012].

Para o desenvolvimento deste trabalho, foram estudadas 15 ferramentas de auxílio na verificação e detecção de indícios de plágio, com o intuito de levantamento de algumas características, das quais foram observadas:

- Tipo de licença, livre ou paga;
- Formato de apresentação e execução da ferramenta para o usuário, Web ou Desktop;
- Tipos de extensões de arquivos analisados;
- Tamanhos de arquivos suportados para análise;
- Necessidade de cadastro para utilização;
- Possibilidade de integração com ambientes virtuais de aprendizagem;
- Geração e apresentação relatório dos resultados da análise detalhando os indícios de plágio.

As ferramentas estudadas foram: Araponga [SANTOS e FRANCO, 2010], DIP – Detector de Índícios de Plágio [PERTILE, 2011], DOCCOP [DOCCOP, 2012], Ephorus [EPHORUS, 2012], Etblast [ETBLAST, 2012], Farejador de Plágio [FAREJADOR, 2012], Plagiarism Detect [PLAGIARISM.ORG, 2012], Plagiarisma [PLAGIARISMA, 2012], Plagium – Online [PLAGIUM, 2012], Plagius Detector [PLAGIUS, 2012], ScholarOne [SCHOLARONE, 2012], Turnitin [TURNITIN, 2012], Urkund [URKUND, 2012], VIPER [VIPER, 2012], todas serão descritas nas próximas seções.

Abaixo, na Tabela 1, encontra-se um breve comparativo das ferramentas estudadas.

Após realizar o levantamento dessas ferramentas que foram listadas (Tabela 1), consideraram-se as características apresentadas e observou-se o quanto ainda se faz necessário a melhoria e implementação de novas ferramentas

Tabela 1. Ferramentas para detecção de indícios de plágio

Ferramenta	Gratuita	Paga	WEB	Desktop	Extensões suportadas	Cadastro	Integrada a AVAs	Apresenta relatório
Araponga [SANTOS e FRANCO, 2010]	Sim	Não	Sim	Não	∅	Sim	Sim	Sim
DIP – Detector de Indícios de Plágio [PERTILE e MEDINA, 2011]	Sim	Não	Sim	Sim	.doc	Não	Sim	Sim
DOCCOP [DOCCOP, 2012]	Sim	Não	Sim	Não	.doc e .pdf	Não	∅	Sim
Ephorus [EPHORUS, 2012]	Não	Sim	Sim	∅	∅	Sim	Sim	Sim
Etblast [ETBLAST, 2012]	Sim	Não	Sim	Não	∅	∅	∅	Sim
Farejador de Plágio [FAREJADOR, 2012]	Sim	Sim	Não	Sim	.doc e .rtf	Sim	∅	Sim
Plagiarism Detect [PLAGIARISM.ORG, 2012]	Sim	Sim	Sim	Sim	∅	Sim	∅	Sim
Plagiarisma [PLAGIARISMA, 2012]	Sim	Não	Sim	Não	∅	Não	∅	Sim
Plagium – Online [PLAGIUM, 2012]	Sim	Não	Sim	Não	.txt	Não	Não	Sim
Plagius Detector [PLAGIUS, 2012]	Sim	Sim	Não	Sim	.doc, .pdf, .rtf, .HTML, .txt	Sim	∅	Sim
ScholarOne [SCHOLARONE, 2012]	Não	Sim	∅	∅	∅	∅	∅	Sim
Turnitin [TURNITIN, 2012]	Não	Sim	Sim	Sim	∅	Não	Sim	Sim
Urkund [URKUND, 2012]	Não	Sim	Sim	∅	∅	∅	Sim	Sim
VIPER [VIPER, 2012]	Sim	Não	Sim	Sim	∅	Sim	∅	Sim

Legenda: ∅ = não foram encontradas informações.

para a detecção de indícios de plágio, pois ainda apresentam limitações, conforme as observadas e listadas acima. Sendo assim, projetou-se a ferramenta Miss Marple.

### 3. MÉTODO DIP – DETECTOR DE INDÍCIOS DE PLÁGIO

O Método DIP trata-se de uma ferramenta

desenvolvida em [PERTILE, 2011], que analisa e calcula o percentual de indícios de plágio. São verificados arquivos com extensão .doc através de pesquisas realizadas na Internet, utilizando a API de buscas do Google - API Google Search Ajax. As pesquisas são realizadas a partir do envio de sentenças contendo no máximo cinquenta palavras, e essas sentenças são compostas por palavras que contenham significado, ou seja, palavras que não

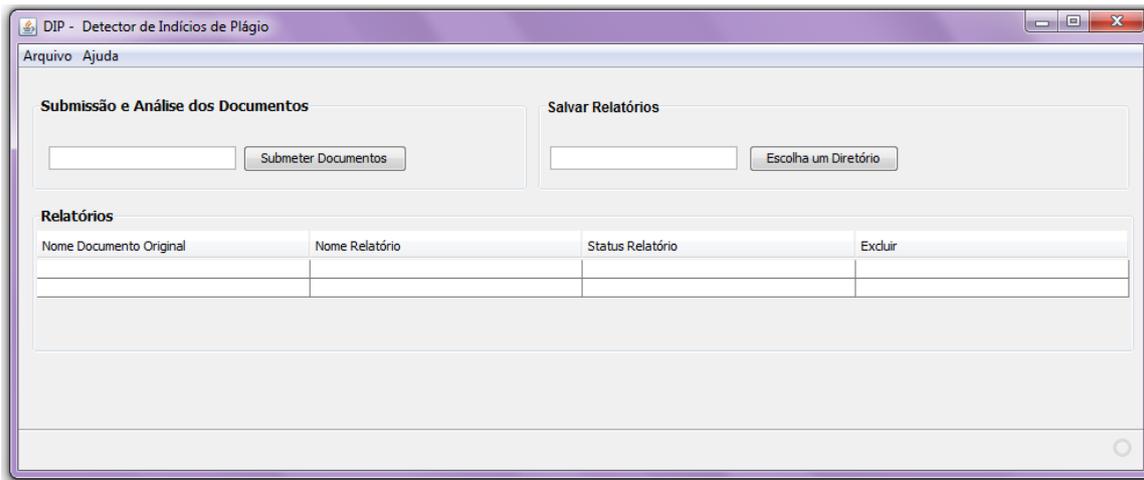


Figura 1 - DIP – Versão Desktop. [PERTILE, 2011]

são preposições e/ou artigos, bem como é realizado a remoção de espaços em branco e figuras. A esta etapa dá-se o nome de remoção de stopwords ou pré-processamento textual. Após esta etapa, realiza-se em seguida o cálculo da similaridade e o percentual de indícios de plágio, que se utilizam dos termos enviados para pesquisa pela API na Web, a constatação de similaridade é feito a partir da comparação dos termos enviados e o resumos das pesquisas retornadas – content -.

Os termos enviados pela API são comparados em relação ao content da pesquisa e então calcula-se a similaridade e o percentual de indícios de plágio, o resultado desses cálculos é apresentado no relatório ao final da análise de

cada arquivo submetido. DIP apresenta-se em duas versões, desktop e acoplada ao AVA Moodle e MLE Moodle.

O funcionamento do DIP ocorre da seguinte forma: inicialmente, o usuário escolhe um diretório onde deseja salvar seus relatórios retornados pelo DIP. Em seguida, os arquivos são submetidos no DIP, que envia parágrafos com um número x de palavras para análise, a análise se dá a partir da comparação dos parágrafos com o content das pesquisas do Google.

Após é gerado um relatório, em pdf, no qual é apresentado o percentual de indícios de plágio de cada parágrafo bem como as URLs onde se encontram. Nas Figuras 1, 2 e 3, ilustra-se o



Figura 2 - DIP – Versão Moodle. [PERTILE, 2011]



Figura 3 - DIP – Versão MLE-Moodle. [PERTILE, 2011]

DIP, versão desktop, integrado dentro do Moodle e também no MLE-Moodle, respectivamente. [PERTILE, 2011]

#### 4.A FERRAMENTA DESENVOLVIDA: MISS MARPLE

A ferramenta que está sendo proposta utiliza como base de desenvolvimento o Método DIP [PERTILE, 2011]. O Método DIP apresenta algumas restrições como por exemplo: faz análise de arquivos somente com extensões .doc, a análise e o cálculo de similaridade são feitos a partir das buscas feitas na internet sem a comparação de arquivos inteiros, não utiliza de técnicas de stemming que visam a distinção entre palavras que possuem o mesmo radical (neste capítulo, nos parágrafos a seguir será melhor detalhado essa técnica). Com

vista nesses requisitos, é o que se almejou com o desenvolvimento desse trabalho, o qual busca o desenvolvimento de uma nova ferramenta de detecção de indícios de plágio, denominada Miss Marple, para auxiliar na verificação da autenticidade dos trabalhos acadêmicos e ou textos publicáveis. A Tabela 2 resume as diferenças entre o Método DIP e a ferramenta Miss Marple.

Na ferramenta Miss Marple, o usuário submete arquivos com diversas extensões, como .doc, .docx, .pdf, ou rtf. Todos os arquivos, independente do tipo de extensão utilizada, são convertidos em texto puro .txt, durante o pré-processamento onde ocorre a remoção das figuras, espaços, remoção das stopwords (palavras que são consideradas irrelevantes na análise de indícios de plágio, por exemplo: advérbios, artigos, conjunções, preposições e pronomes [DIAS, 2004]) e indexação de palavras, por fim, são enviados sentenças compostas por 50 palavras para a API de pesquisa do Google - API Google Search Ajax. Esta API examinará a similaridade com o content (breve descrição do que se trata a pesquisa) do Google e coletará as fontes/URLs, das quais serão feitas o download dos arquivos que contenham pelo menos 60% de termos similares em seu content. O percentual limite para demonstração e análise de indícios de plágio foi fixado em 60% devido os testes já realizados em [PERTILE, 2011] que mostrou que esse percentual foi o que melhor apresentou resultados relevantes.

Após o download, esses arquivos são armazenados em um diretório na máquina do usuário que formará o repositório dos arquivos para análise detalhada. Assim, é nesta biblioteca digital que será realizada a avaliação de referências cruzadas, ou seja, será realizada a análise de indícios de plágio do documento original em relação aos documentos que fazem parte do repositório. Na análise, também será levado em consideração a indexação dos ter-

Tabela 2. Resumo das diferenças entre Método DIP e ferramenta Miss Marple

	Extensão de arquivos	Análise	Utilização de Técnicas de stemming
Método DIP – Detector de Índicios de Plágio	apenas .doc	- Feitas a partir da API de buscas do Google - <i>API Google Search Ajax</i> , onde os arquivos são enviados para análise e a comparação e o cálculo de similaridade é feito em relação ao <i>content</i> retornados da pesquisa	Não
Miss Marple – Detector de Índicios de Plágio	.doc, .docx, .pdf, ou HTML	- Feitas a partir da API de buscas do Google - <i>API Google Search Ajax</i> - Análise em relação ao arquivos que compõe o repositório	Sim

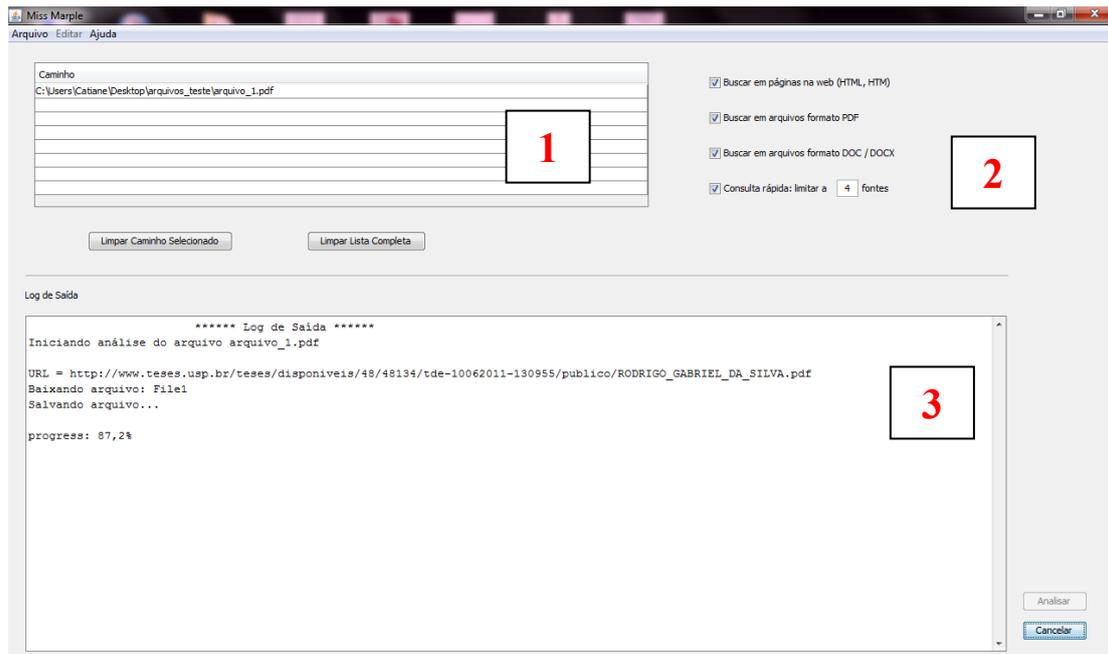


Figura 4 – Interface do Miss Marple, iniciando a execução de uma análise

mos a partir da utilização da biblioteca Lucene – Java, essa biblioteca possibilita a aplicação de técnicas de stemming ao indexar termos para que assim facilite a comparação de termos que contenham o mesmo radical, por exemplo: Carro e Carroça, ambos têm o mesmo radical, mas são palavras distintas. Ao final da análise na Internet e análise no repositório, é gerado um relatório que apresenta o percentual de indícios de plágio bem como as referências nas quais foram encontrados os indícios e as URLs (endereço eletrônico onde se localizam os textos similares).

Na Figura 4, acima é ilustrado a ferramenta Miss Marple iniciando a execução de uma análise.

Foram marcados na Figura 4, alguns itens que serão descritos:

- Ítem 1: Nesta area fica localizada a lista de arquivos que estão sendo analisados em um determinado instante;
- Ítem 2: Localiza-se os controles de tipos de buscas que o usuário deseja fazer (tipos de arquivos para compor o repositório), possibilidade de limitar o número de fontes pesquisadas;
- Ítem 3: No decorrer da execução da análise, é dado um feedback em tempo real das ações que estão ocorrendo no processo de análise. Neste espaço são descritos: início da análise de cada arquivo, formação do repositório, gravação dos arquivos no repositório, final da análise, tempo decorrido no processo de análise, endereços dos arquivos baixados, e score com percentual de indícios de plágio e similaridade de cada arquivo em relação ao submetido.

Na Figura 5, é ilustrado a mensagem de fim de uma análise.

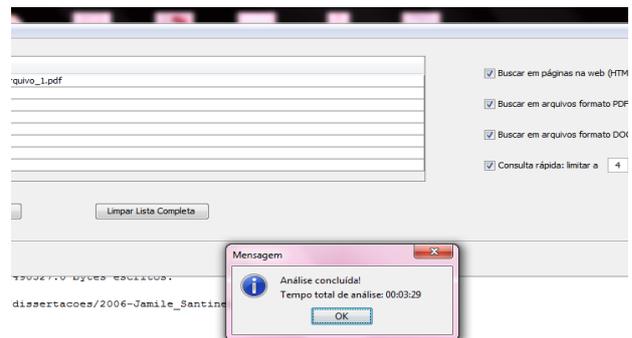


Figura 5- Análise de um texto concluída

A seguir, na Figura 6 é ilustrado o feedback de saída do final da análise. Neste feedback, ilustrado na figura 6, as setas apontam os endereços de cada um dos arquivos que foram impressos no relatório final, bem como o score com o percentual de similaridade em relação ao arquivo submetido. Além da frase ao final que destaca a conclusão da análise.

Em seguida, na Figura 7 é apresentada a formulação do repositório da análise executada.

Neste repositório, conforme visto na Figura 7, encontram-se os arquivos que foram utilizados para análise de indícios de plágio, além dos arquivos, encontra-se também um relatório com o endereço dos arquivos na WEB, localização no computador e score com percentual de similaridade entre o arquivo submetido e o baixado.

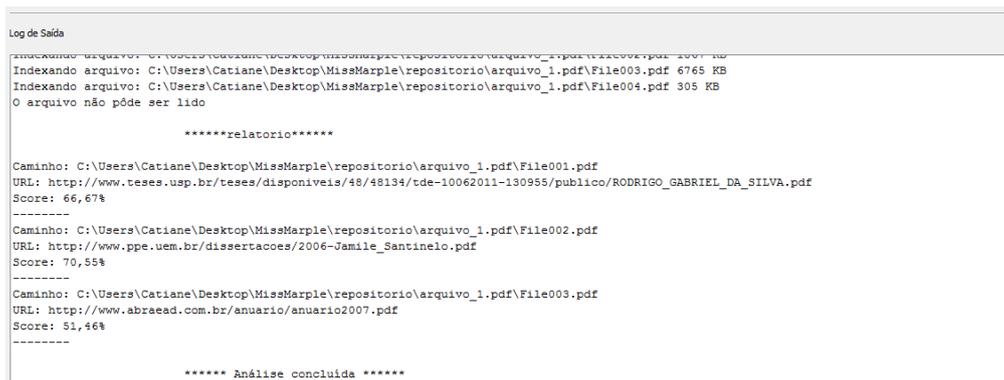


Figura 6- Feedback de final da análise

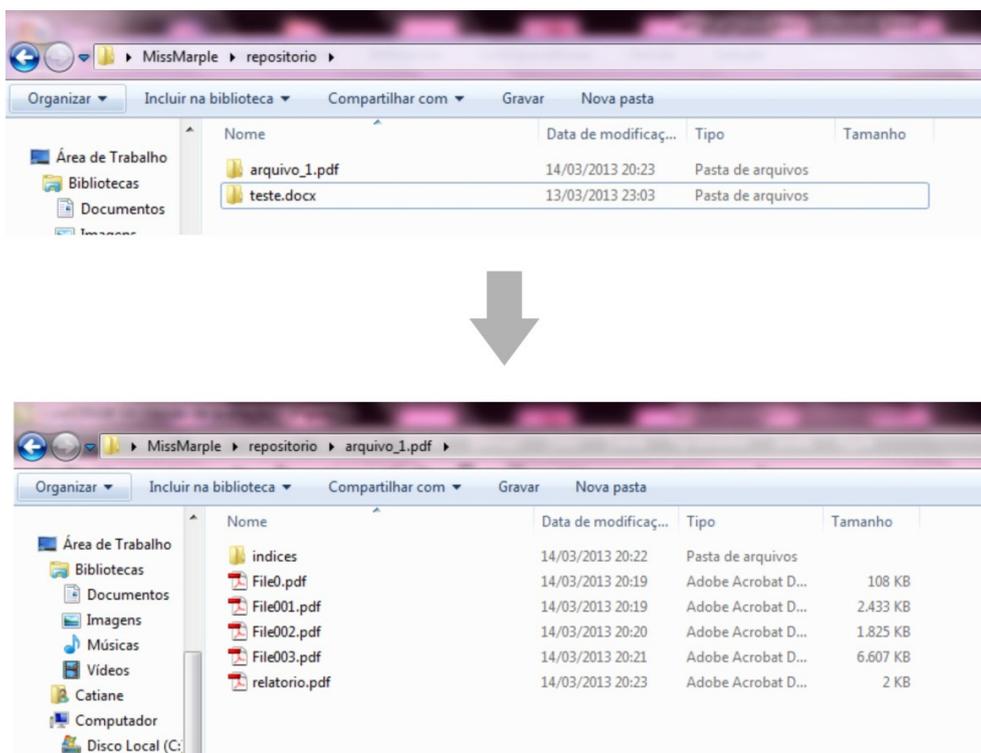


Figura 7 – Formação do repositório.

## 5 RESULTADOS

A metodologia de testes utilizada para a validação da ferramenta consiste na seleção de oito artigos desenvolvidos por alunos de um curso de nível de especialização de uma universidade “A” Federal, oito artigos de um curso de nível de graduação de uma universidade “B” privada e dois textos escritos sem qualquer indício de plágio para fins de confirmação de autenticidade do score apresentado como percentual de indícios de plágio. O total selecionado de artigos não foi utilizado, apenas alguns fizeram parte dos testes. Todos os textos submetidos para análise tinham entre 8 e 15 páginas e foram escolhidos aleatoriamente

para os testes.

Para tal análise, foram elencadas três ferramentas das quais foram estudadas no decorrer do trabalho além da ferramenta desenvolvida. As ferramentas escolhidas foram: VIPER, Farejador de Plágio, Plagius Detector e a ferramenta desenvolvida nesse trabalho denominada Miss Marple.

Nos testes realizados, foram levados em consideração: tempo de processamento, atendimento de requisitos de usabilidade de software através de um questionário adaptado da proposta de checklist de [NUNES et. al, 2012] e precisão dos resultados que também foi abordado no quesito usabilidade.

### 5.1 Tempo de processamento

Em ambas as ferramentas foram submetidos os mesmos arquivos, utilizando para esses testes um Notebook (Sistema Microsoft Windows 7 - Service Pack 3 - Intel (R) Core 2 Duo (R), 4Gb de RAM), internet via rádio 3Mb e o mesmo hardware com internet 512Kb. Todos os testes foram realizados em um mesmo horário, porém em dias distintos. No Gráfico 2, é ilustrado o tempo de processamento da Internet com velocidade de 3Mb.

Neste Gráfico 2, encontram-se os dados de análise de quatro arquivos do total de 18 que foram selecionados para verificação. Estes arquivos foram selecionados aleatoriamente do total disponível para verificação.

A partir do momento da seleção, os mesmos arquivos foram submetidos na diferentes

ferramentas, os arquivos foram submetidos em dias diferentes e no mesmo horário de acesso a Internet para fins de precisar o tráfego da rede no dado instante. O Gráfico 2 apresenta a ferramenta Miss Marple em destaque em relação ao tempo de processamento de análise e apresentação do resultado final. O diferencial da ferramenta está na possibilidade de o usuário limitar sua busca a um número x de fontes, dependendo de sua velocidade de Internet e disponibilidade de tempo. Neste caso, (Gráfico 2) o limite de buscas por arquivos suspeitos foram fixados em 10 fontes de pesquisa para análise e comparação, e todas as ferramentas têm como filtro realizar a busca e a amostragem das 10 fontes mais relevantes ou que apresentam maior porcentagem de semelhança com o arquivo submetido para análise.

O Gráfico 3 traz os dados dos testes feitos

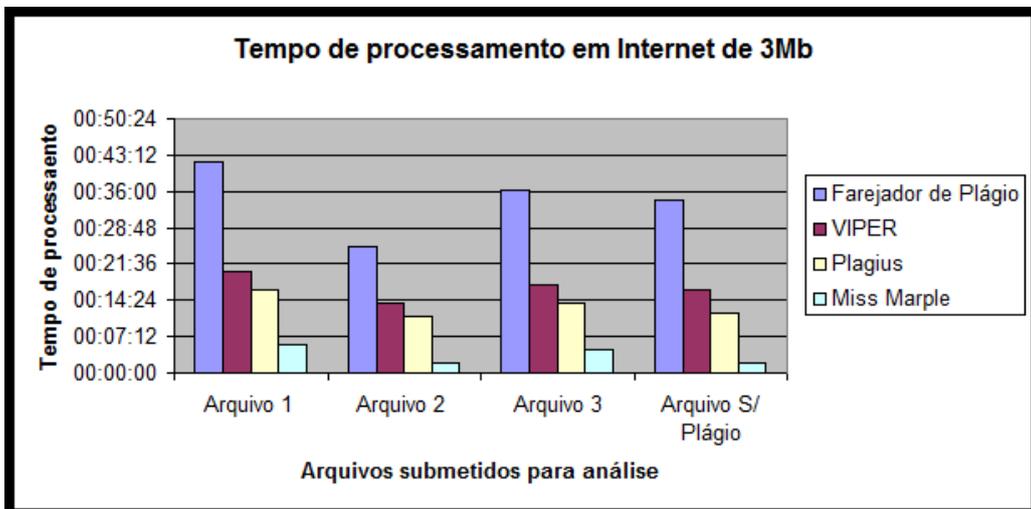


Gráfico 2 - Tempo de processamento em Internet de 3Mb

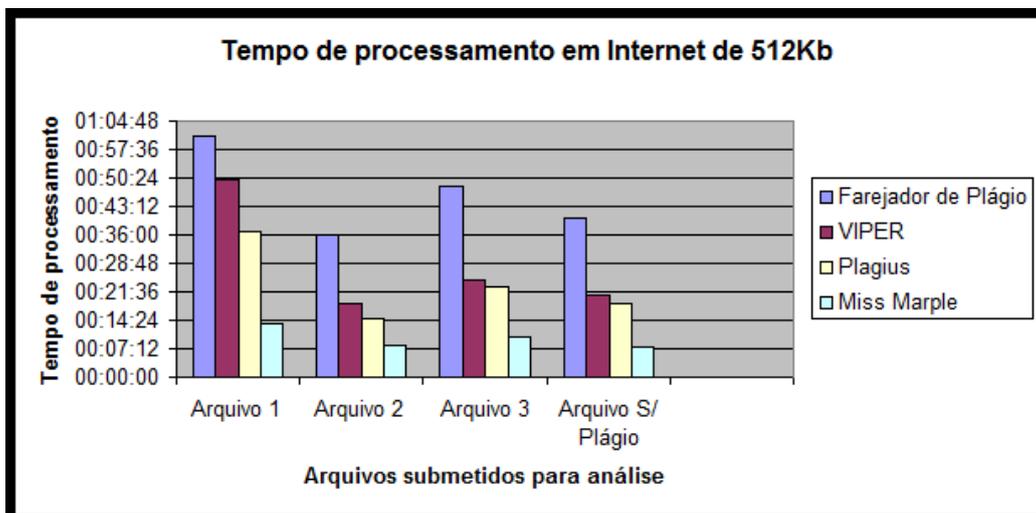


Gráfico 3 - Tempo de processamento em Internet de 512Kb

em uma velocidade de 512 Kb, com limite de buscas em 10 fontes.

A metodologia adotada para análise e coleta de amostragem dos dados foi a mesma adotada nas análises ilustradas no Gráfico 3. A ferramenta com melhor tempo de processamento segue sendo a Miss Marple, seguida pela Plagius e VIPER, e a ferramenta que apresentou em todos os testes maior tempo de processamento foi a ferramenta Farejador de Plágio. No Gráfico 3 também é possível identificar o quanto a velocidade de Internet contratada é um fator diferencial no tempo de processamento, já que ambas as ferramentas disponíveis no mercado e usadas nas análises fazem a busca por textos suspeitos na Web.

## 5.2 Outros aspectos analisados

No trabalho, foram analisados outros aspectos, tais como ocorrência de falhas na execução, consistência das referências encontradas e alguns quesitos de funcionalidades, baseados na proposta de [NUNES et. AL, 2012].

Esses aspectos analisados são ilustrados nos Gráficos 4, 5 e 6 abaixo.

O Gráfico 4 aborda o questionamento sobre a frequência de ocorrência de falhas durante a utilização da Ferramenta Miss Marple

Os resultados apresentados no gráfico acima são satisfatórios: 73% dos usuários responderam que a ferramenta desenvolvida não apresenta falhas durante a execução, 6% responderam que a ferramenta apresenta falhas no momento da exclusão

do diretório criado após a criação do repositório e 21% consideram que as falhas encontradas podem ser parciais. Nenhum usuário encontrou falhas que comprometessem a qualidade da análise de ocorrência de indícios de plágio.

O Gráfico 5 traz os dados referentes a consistência das referências/arquivos apontados como possíveis indícios de plágio, ou seja, há semelhanças entre o arquivo submetido e o encontrado na Web.

Do total de entrevistados, 84% avaliou que as referências encontradas com semelhanças em relação ao arquivo submetido para análise, ou com indícios de plágio, são referências consistentes ou corretas. 16% dos entrevistados apontaram como parciais.

O Gráfico 6 apresenta quesitos relacionados à facilidade de utilização da Ferramenta Miss Marple.

Nos questionamentos contemplados e apresentados no Gráfico 6, os usuários avaliaram a facilidade de uso da ferramenta, e 100% dos participantes responderam que é fácil a sua utilização. Em relação à facilidade de aprendizagem de utilização do software, 92% dos usuários responderam que a ferramenta é fácil de aprender e 8% respondeu que é parcialmente fácil. A ferramenta dispõe de manual de ajuda e botão de ajuda dentro dela mesma e esse quesito foi avaliado e confirmado por 100% dos usuários. Em relação à apresentação de feedback, manter os usuários informados da ocorrência dos processos de análise, 90% dos usuários responderam que o feedback é satisfatório e 10% respondeu que é parcialmente satisfatório.

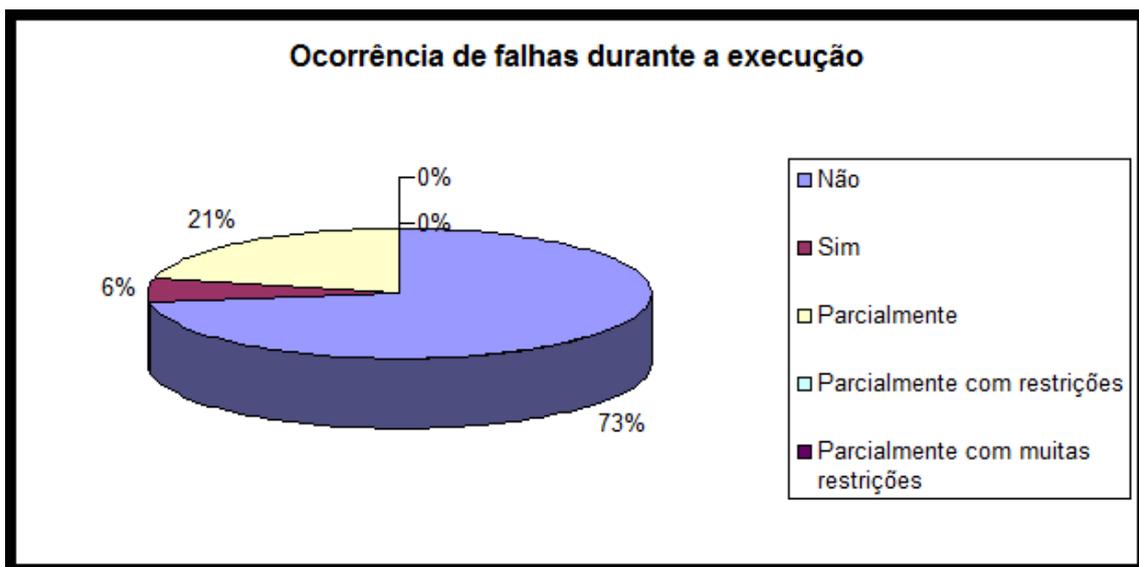


Gráfico 4 – Ocorrência de falhas durante a execução do Miss Marple

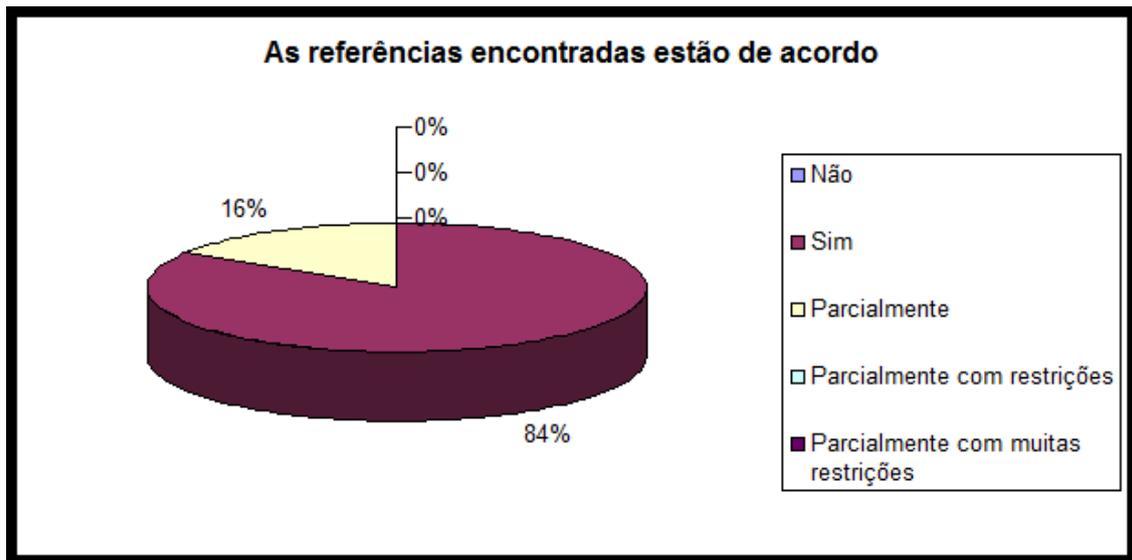


Gráfico 5 – Consistência das referências encontradas

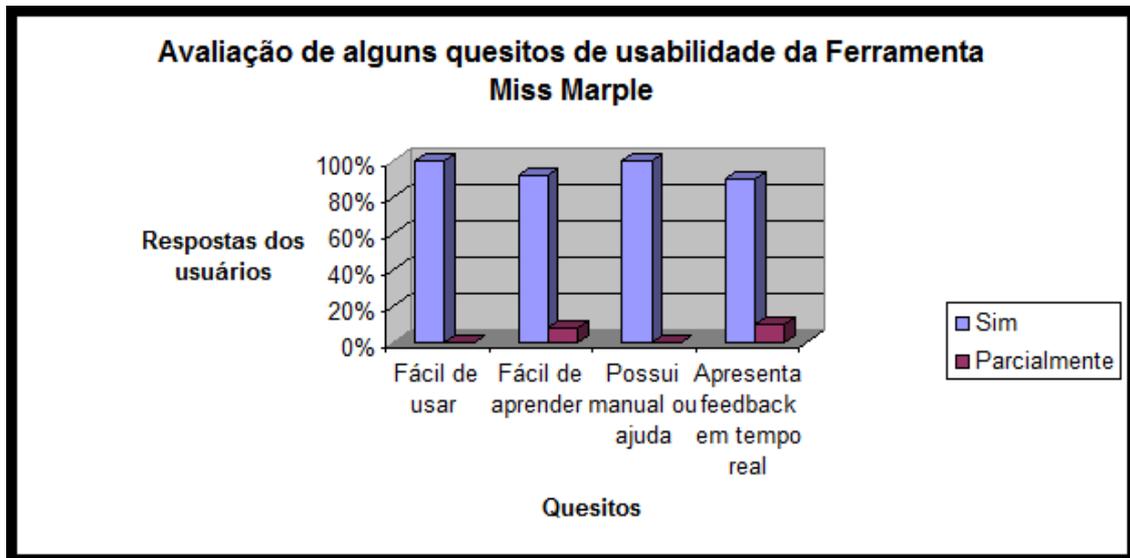


Gráfico 6 – Alguns dos quesitos avaliados no checklist de usabilidade.

## 6. CONSIDERAÇÕES FINAIS

A evolução da computação e consequentemente da velocidade e quantidade de informações disponíveis acabou por remeter os olhos dos professores para a verificação de autenticidade dos trabalhos desenvolvidos pelos alunos. Por se tratar de uma tarefa dispendiosa de tempo é importante contar com o auxílio de softwares. Durante as pesquisas, foram encontradas várias ferramentas com a finalidade de análise de indícios de plágio, porém poucas de uso gratuito, e todas apresentavam algum tipo de restrição. Contudo, verificou-se o

quanto ainda são necessários o aprimoramento e desenvolvimento de novas ferramentas, desta maneira, é que foi proposto o desenvolvimento da ferramenta Miss Marple para auxiliar na análise de indícios de plágio.

Enfim, a ferramenta desenvolvida e testada apresentou resultados satisfatórios conforme os resultados que foram apresentados.

## 7. REFERÊNCIAS

- BARNBAUM, C. PLAGIARISM: A Student's Guide to Recognizing It and Avoiding It. Valdosta State University, (2002). Disponível em: <[http://www.valdosta.edu/~cbarnbau/personal/teaching\\_MISC/plagiarism.htm](http://www.valdosta.edu/~cbarnbau/personal/teaching_MISC/plagiarism.htm)>. Acesso em: 15 de julho de 2012.
- DIAS, M. A. L. Extração Automática de Palavras-Chave na Língua Portuguesa Aplicada a Dissertações e Teses da Área das Engenharias, 2004. 127 f. Dissertação (Mestrado em Engenharia Elétrica) - Faculdade de Engenharia Elétrica e de Computação, Campinas, SP.
- DOC COP. (2012). Disponível em: <<http://www.doccop.com/>>. Acesso em: 20 de junho de 2012.
- EPHORUS. Ephorus: liderança na Europa, (2012). Disponível em: <<http://www.ephorus.pt/home>>. Acesso em: 21 de julho de 2012.
- ETBLAST, 2012. Disponível em: <<http://etest.vbi.vt.edu/etblast3/>>. Acesso em: 20 de junho de 2012.
- FAREJADOR. Farejador de Plágios, (2012). Disponível em: <<http://www.farejadordeplagio.com.br/>>. Acesso em: 21 de julho de 2012.
- FURTADO, J. A. X. P. Trabalhos acadêmicos em Direito e a violação de direitos autorais através de plágio Disponível em: <<http://www.egov.ufsc.br/portal/sites/default/files/anexos/5640-5632-1-PB.htm>> Acesso em: 21 de julho de 2012.
- HANDBOOK, A. (Brasil). 07.11 - Code of Practice on Plagiarism, v. 1, (2009).
- LIMA, C. E., RESENDE, P. M. A., 2012. Análise qualitativa e quantitativa entre as principais ferramentas de detecção de plágio. Disponível em: <[http://www.c3.furg.br/arquivos/download/04\\_lima\\_resende.pdf](http://www.c3.furg.br/arquivos/download/04_lima_resende.pdf)> Acesso em 15 de agosto de 2012.
- MORAES, R. O plágio na pesquisa acadêmica: a proliferação da desonestidade intelectual. In: Revista Diálogos Possíveis - Faculdade Social da Bahia, Bahia, n. 1, p. 92-109, jun. 2004. Disponível em: <<http://www.faculdadesocial.edu.br/dialogospossiveis/artigos/4>>. Acesso em: 02 de julho de 2012
- NEIL, R. (2004). Cheating in online student assessment: Beyond plagiarism. Online Journal of Distance Learning Administration, Volume VII, Number II, State University of West Georgia, Distance Education Center.
- NUNES, F. B. ; VOSS, G. B. ; MUHLBEIER, A. R. K. ; MEDINA, R. D. ; BERNARDI, G. ; BARBOSA, C. P. A. . Análise Comparativa Teórico-Prática entre Softwares de Detecção de Plágio. RENOTE. Revista Novas Tecnologias na Educação, v. 10, p. 1-10,
- OLIVEIRA, M. et al.. Bibliotecas Digitais Aliadas na Detecção Automática de Plágio. Seminário Internacional de Bibliotecas Digitais Brasil. 2007. Disponível em: <<http://libdigi.unicamp.br/document/?code=23482>>. Acesso em: 05 de julho de 2012.
- OLIVEIRA, M. G. D.; OLIVEIRA, E. Uma Metodologia para Detecção Automática de Plágios em Ambientes de Educação a Distância. In: Congresso Brasileiro de Ensino Superior a Distância – ESUD 2008, Gramado, RS, 2008. 1-20.
- PLAGIARISMA. (2012). Disponível em: <<http://plagiarisma.net/>>. Acesso em: 20 de junho de 2012.
- PLAGIARISM.ORG, 2012. What is plagiarism? Plagiarism.org. Disponível em: <[http://www.plagiarism.org/plag\\_article\\_what\\_is\\_plagiarism.html](http://www.plagiarism.org/plag_article_what_is_plagiarism.html)>. Acesso em: 21 de julho de 2012.
- PLAGIUM. (2012). Disponível em: <<http://www.plagium.com/>>. Acesso em: 12 julho de 2012.
- PLAGIUS. Plagius - The ultimate in plagiarism detection, 2012. Disponível em: <<http://www.plagius.com/en/default.aspx>>. Acesso em: 21 de julho de 2012.
- PLAGIO.NET, 2012. Disponível em: <<http://www.plagio.net.br/pesquisa-e-publicacoes.html>>. Acesso em: 02 de julho de 2012
- PERTILE, S. L. ; MEDINA, R. D. . Desenvolvimento e Aplicação de um Método para Detecção de Indícios de Plágio. In: Simpósio Brasileiro de Informática na Educação, 2011, Aracajú. Anais do XXII SBIE - XVII WI, (2011). p. 1673-1682.
- . Desenvolvimento e Aplicação de um Método para Detecção de Indícios de Plágio. In: Conferência IADIS Ibero Americana WWW/INTERNET 2011, 2011, Rio de Janeiro. Conferência IADIS Ibero Americana WWW/INTERNET (2011).
- . “Desenvolvimento e Aplicação de um Método para Detecção de Indícios de Plágio”. Dissertação apresentada ao Curso de Mestrado do Programa de Pós-Graduação em Informática, Universidade Federal

de Santa Maria (UFSM, RS), 2011.

SANTOS, A. O. F., FRANCO, R. H. R. L. (2010) Criação de Ferramenta de Detecção de Plágio em Ambiente Virtual de Aprendizagem. Dissertação apresentada ao Curso de Mestrado do Programa de Pós-graduação em Engenharia Elétrica, Universidade Federal de Itajubá-MG. Disponível em: < <http://adm-net-a.unifei.edu.br/phl/pdf/0037064.pdf>>. Acesso em: 02 de julho de 2012

SIBI. Sistema Integrado de Bibliotecas - Universidade de São Paulo. (2011)

Disponível em: < [http://www.workshop.sibi.usp.br/relatorios/Lista\\_softwares\\_prevencao\\_plagio.pdf](http://www.workshop.sibi.usp.br/relatorios/Lista_softwares_prevencao_plagio.pdf)>

Acesso em: 01 de agosto de 2012.

SCHOLARONE. (2012). Disponível em: < <http://scholarone.com/>>. Acesso em: 21 de julho de 2012.

TURNITIN. Prevent plagiarism, (2012). Disponível em: <<http://turnitin.com/static/index.html>>. Acesso em: 25 de julho de 2012.

URKUND. (2012). Disponível em: <<http://www.urkund.com/int/en/>>. Acesso em: 25 de julho de 2012.

VIPER. The Anti-plagiarism Scanner, (2012). Disponível em: <<http://www.scanmyessay.com>>. Acesso em: 25 de julho de 2012.