
Sumarização de Texto em Ambientes Educacionais na Web

Fernanda Josirene de Melo Ferreira¹, Hemilis Joyse Barbosa Rocha², Gilberto Pedro da Silva Júnior², Ig Ibert Bittencourt¹, Evandro de Barros Costa¹.

¹Instituto de Computação – Universidade Federal de Alagoas (UFAL)
Campus A. C. Simões, BR 104 - Norte, Km 97, Cid. Universitária, Maceió, AL – Brasil

²Faculdade de Alagoas (FAL)
Rua Pio XII, 70, Jatiúca, Maceió/AL - 57035-560, Maceió, AL – Brasil
{fynanda, joyseel, gp7junior, ig.ibert, ebcosta}@gmail.com

Abstract. *As we can not ensure that automatic summarization is viable in the Educational Environment on the web, this motivated us make this research. Due to documents congestion in these environments, a lot of time is wasted by the users when they read a lot of information which are not relevant. For that reason our objective is to explore the usage of automatic summarization and based in their informativity and quality level evaluate how close it is from human's summarization. The results of this study will help determining whether its insertion in these environments is feasible or not, enabling the user to access information faster and simplifying decision making.*

Keywords: *Automatic Summarization, Educational Environment on the Web, Informativity, Quality.*

Resumo. *Como ainda não se pode garantir que a sumarização automática seja viável em Ambientes Educacionais na Web, surge a motivação para a realização desta pesquisa. Devido ao acúmulo de documentos nesses Ambientes, há uma perda de tempo dos usuários quando lêem informações irrelevantes ao interesse. Por esse motivo, objetiva-se explorar os cenários de utilização dessa sumarização e com base no seu nível de informatividade e qualidade, avaliar o quanto ela está próxima da humana. Esses resultados contribuirão para analisar se a sua inserção nesses Ambientes é viável, possibilitando ao usuário o acesso com mais rapidez à informação e facilitando uma tomada de decisão.*

Palavras-Chave: *Sumarização Automática, Ambiente Educacional na Web, Informatividade, Qualidade.*

1. Introdução

A sumarização textual está presente em nosso dia-a-dia sem que possamos perceber, a exemplo disso na escolha de um filme ou livro através das resenhas e até mesmo nos resumos de artigos científicos. Segundo Mani, Bloedorn e Gates (2007, apud Marques Neto, Nascimento e Gomes, p. 01), sumarizar é o processo de seleção das informações mais importantes de um texto para produzir uma versão mais curta. Os sumários ajudam

a identificar de maneira mais rápida, para o leitor, o conteúdo do texto-fonte e assim o mesmo irá distinguir quais materiais são de seu interesse ou não.

Com o crescimento da rede mundial de computadores, tem-se percebido a aglomeração desmedida de documentos na Internet. Não obstante, é nos Ambientes Educacionais na Web que encontramos esta mesma tendência. Além do mais, para Fuks, Gerosa e Pimentel (2003, p. 08), o texto escrito é a linguagem predominante das ferramentas de comunicação para dar suporte ao trabalho em grupo. A necessidade de consumir tais informações, em grande quantidade de maneira completa e no menor tempo possível, torna-se indispensável automatizar o processo de sumarização de textos. Estejam esses em forma de arquivo, de chat ou de blog.

Conforme Oliveira (2008, p. 4) sumarização automática é uma técnica na qual, o computador simula a atividade humana na elaboração de sumários para documentos de texto. Esta vem sendo estudada com bastante ênfase nos últimos anos, com intuito de facilitar as buscas na internet. Assim, acelerou-se o desenvolvimento de aplicativos que sumarizassem a diversidade de informações nela contidas e retornassem informações mais precisas já sumarizadas.

Como ainda não se pode garantir que a sumarização automática seja viável em Ambientes Educacionais na Web, surge uma motivação para avaliarmos tal viabilidade.

Este artigo tem como objetivo explorar os cenários de utilização da sumarização automática nas ferramentas: repositório de arquivos textuais, blog e bate papo dentro do Ambiente Educacional na Web. Além de analisar o nível de informatividade e qualidade dessas sumarizações, para assim avaliarmos se a automática está próxima da humana. Com isso, diante dos resultados da avaliação, verificar se a inserção da sumarização automática no Ambiente Educacional na Web é viável. Dessa forma, facilitará à tomada de decisão e o acesso rápido à informação por parte dos usuários ao acessar a ferramenta para auxiliar no processo de ensino aprendizagem.

O presente artigo está dividido em 6 seções. Na próxima seção serão abordados alguns trabalhos relacionados. Na seção 3, os cenários textuais usados no Ambiente Educacional na Web. Em seguida, na 4ª seção, a metodologia utilizada na avaliação da sumarização textual, na 5ª os resultados dessa avaliação. E por fim, na última seção, será apresentada a conclusão, por meio das considerações finais e sugestões para trabalhos futuros.

2. Trabalhos Relacionados

Com o intuito de entender e levantar os trabalhos existentes que avaliam a qualidade da sumarização automática foi analisado algumas iniciativas. Dessa forma, explanaremos algumas perspectivas do desenvolvimento de trabalhos relacionados.

O foco do trabalho de Gonçalves (2008) está relacionado à qualidade dos sumários extrativos. Por muitas vezes as frases selecionadas no texto não preservam a coesão referencial necessária para o entendimento do texto. Assim foi desenvolvido um sistema que realiza a manutenção da coesão referencial dos sumários extrativos usando como fonte de informação as cadeias de correferência presentes no texto-fonte. A autora ainda fez experimentos e avaliação dos resultados através de 2 sumarizadores: Gist-Summ e Supor-2. No entanto, podemos constatar que o questionário avalia apenas a

legibilidade e a informatividade das sumarizações, diferente do nosso estudo, que ainda contribui com a análise das palavras-chaves.

Diante da sobrecarga dos repositórios textuais, e a dificuldade das pessoas em tomar decisões em curto espaço de tempo, a dissertação de Oliveira (2008) propõe que um juiz humano avalie as sumarizações automáticas geradas por quatro ferramentas. Dessa forma ele aborda um comparativo das quatro em relação a informatividade e a textualidade. Enquanto que em nosso trabalho, compara-se a sumarização automática com a sumarização humana, em relação ao quanto elas se aproximam.

A dissertação de Coracini (2009) apresenta estratégias atualmente em uso para uma avaliação sistemática, assim como as necessidades e limitações dessa tarefa. Tentando-se descobrir como o nível de compreensão textual, pois não se tem acesso as representações e operações mentais relacionadas no processo. Com uma pesquisa exploratória a autora busca avaliar o quanto o sumário ajuda na compreensão de textos através da análise de alunos universitários. Como foi citado anteriormente esse trabalho não se preocupa em avaliar as palavras-chave, nem comparar a sumarização automática com a sumarização manual.

O objetivo do trabalho dos autores Balage Filho, Pardo e Nunes (2007) é aprimorar o método de sumarização do sistema GistSumm para sanar algumas de suas limitações observadas, como o tratamento de textos estruturados, que possuem seções e subseções. Sendo assim este trabalho apresenta uma avaliação da sumariizador GistSumm para textos científicos, versus o GistSumm original. Para isso duas avaliações distintas foram conduzidas: avaliação baseada no julgamento humano e avaliação baseada no julgamento da ferramenta ROUGE para avaliação automática de sumários. No entanto, os autores não confrontam os resultados das avaliações, enquanto que nossa pesquisa confronta os resultados da avaliação da sumarização humana, versus sumarização automática.

3. Cenários de Uso

Cada item abordado nesta seção corresponde a uma proposta de sumarização para ser empregada dentro de um Ambiente Educacional na Web, dependendo dos resultados obtidos na avaliação descrita na seção 5. Dessa forma, entenderemos o funcionamento de cada cenário a ser usado como apoio no experimento da sumarização.

a) Repositório de Arquivos Textuais

Segundo Bertagnolli et al (2007, p. 12), o repositório de arquivos textuais “é uma ferramenta que permite que os professores disponibilizem material adicional a suas aulas, como por exemplo: uma planilha, um documento do Word, transparências, entre outros”.

Com o passar do tempo, este tipo de ferramenta irá ficar sobrecarregada com o acúmulo de arquivos textuais, e com isso, dificultará o acesso a estes por parte dos usuários do Ambiente Educacional na Web, ao desejar selecionar um assunto de seu interesse. Neste caso, surge 2 cenários.

O primeiro deles é o caso da sumarização estar embutida no repositório, organizando seus dados automaticamente, de forma que quando o usuário for acessar, seja ele professor ou aluno, já esteja disponível o sumário do arquivo a ser acessado.

O segundo cenário é a sumarização personalizada, onde o usuário ao sentir-se insatisfeito com a sumarização pré-construída automaticamente, seja por preferir o sumário menor ou maior, através da definição do nº de palavras-chave, ou do nº de linhas. Ou ainda escolhendo o atributo de visualizar as palavras-chave em destaque ou não, por exemplo. Desta forma, o usuário construirá seu próprio sumário, de acordo com suas preferências em relação aos atributos disponíveis na ferramenta de sumarização do repositório.

b) Blog

Conforme Moresco e Behar (2006, p. 3 e 4), o Blog, “foi criado pela empresa Pyra Labs em 1999, para que as pessoas criassem seus diários na Internet. É uma nova modalidade de leitura e de escrita, um novo meio de produção de conhecimento e de compartilhamento de informações”. Para fins educacionais, o blog pode ser utilizado de forma que o usuário aluno poste diariamente, um relato sobre o que compreendeu de suas investigações durante a realização de determinado projeto. Desta forma, ele será co-autor do processo de aprendizagem.

O usuário uma vez logado no Ambiente Educacional na Web, tem acesso aos perfis dos demais usuários, e conseqüentemente ao Blog dos mesmos. Desta forma, antes de acessar o Blog alheio, o usuário poderá visualizar uma lista com todos os sumários de cada post do Blog, em ordem cronológica. Isto facilitará a busca do post do Blog de acordo com o assunto de interesse do usuário. Assim, o usuário irá decidir se vale a pena ler o post do Blog, antes de lê-lo. Logo, surge o cenário da sumarização do Blog, dentro do Ambiente Educacional na Web.

c) Bate Papo

Para Bertagnolli et al (2007, p. 12), Chat ou bate-papo, “é um ambiente para a troca de mensagens instantâneas, no qual os professores podem disponibilizar a seus alunos várias salas para debate”. Tal ferramenta possibilita o diálogo, a troca de informação e de conhecimento entre pessoas de distintas localidades.

Neste caso, para haver a interatividade nas salas de bate papo, é necessária a comunicação síncrona dentro do Ambiente Educacional na Web, ou seja, a presença simultânea dos participantes envolvidos. Mas, como sabemos, imprevistos acontecem decorrentes da vida diária, e por este motivo pode ocorrer de algum dos usuários, não conseguir se reunir no horário previamente marcado, tendo perdido os assuntos discutidos, novidades ou avisos dados. Apesar dos autores Hernades, Santana e Falcão (2000, p. 8) considerarem vantagem o fato de o bate-papo disponibilizar uma “documentação para auxiliar os alunos que não participaram de uma sessão de chat a se inteirar do que foi discutido”, no entanto torna-se cansativo ler extensivas linhas desnecessárias, e de pouca relevância de tal documentação gravada.

Desta forma, surge o cenário em que o usuário visualiza os links de abertura do bate-papo. Estarão também disponíveis passagens sumarizadas do chat gravado, como

ocorre nas ferramentas de busca, por exemplo, trazendo trechos sumarizados da página a ser acessada.

4. Metodologia

Nessa seção trataremos da metodologia utilizada na comparação da Sumarização Automática com a Sumarização Manual de cada cenário tratado.

Para este trabalho, contamos com dez acadêmicos integrantes do GroW¹ (Grupo de Otimização da Web) do Curso de Ciências da Computação da Universidade Federal de Alagoas. Foram distribuídas entre eles, vinte sumarizações de blogs e vinte conteúdos extraídos da Wikipédia como Repositório de Arquivos Textuais, ou seja, quatro sumários por pessoa para responder o questionário.

O experimento consistiu na criação e aplicação de questões objetivas sobre a sumarização humano-manual versus a sumarização automática. Onde os alunos citados opinaram com respostas que poderiam variar entre “Ótimo”, “Bom”, “Regular”, “Ruim”, “Péssimo”, bem como respostas do tipo “Sim” ou “Não” e ainda escolher entre a sumarização nº 1 ou a nº 2, onde a primeira se refere à sumarização manual, e a segunda se refere à automática. Propositamente, não foi informado aos avaliadores na aplicação do questionário, qual das sumarizações era a manual. A seguir o questionário aplicado:

Tabela 1. Questionário Avaliativo

Sumarização Manual X Sumarização Automática	
Nome:	
Idade:	
1. Como você definiria o nível de qualidade¹ em relação à coerência² das palavras da Sumarização1?	
<input type="checkbox"/> ótima	<input type="checkbox"/> boa <input type="checkbox"/> regular <input type="checkbox"/> ruim <input type="checkbox"/> péssima
2. Como você definiria o nível de informatividade³ da Sumarização1 em relação ao texto original?	
<input type="checkbox"/> ótima	<input type="checkbox"/> boa <input type="checkbox"/> regular <input type="checkbox"/> ruim <input type="checkbox"/> péssima
3. Como você definiria o nível de qualidade em relação à coerência das palavras da Sumarização2?	
<input type="checkbox"/> ótima	<input type="checkbox"/> boa <input type="checkbox"/> regular <input type="checkbox"/> ruim <input type="checkbox"/> péssima
4. Como você definiria o nível de informatividade da Sumarização2 em relação ao texto original?	
<input type="checkbox"/> ótima	<input type="checkbox"/> boa <input type="checkbox"/> regular <input type="checkbox"/> ruim <input type="checkbox"/> péssima
5. Qual você acha que é a sumarização automática?	
<input type="checkbox"/> Sumarização1	<input type="checkbox"/> Sumarização2
6. Você trocaria a Sumarização Manual pela Sumarização Automática?	
<input type="checkbox"/> Sim	<input type="checkbox"/> Não
7. As palavras-chave apresentaram mais relação em qual sumarização?	
<input type="checkbox"/> Sumarização1	<input type="checkbox"/> Sumarização2
¹ O nível de qualidade é se o sumário forma realmente um texto e se tem coerência com as palavras.	
² Coerência, segundo Ferreira (2004), se refere a: harmonia entre idéias e acontecimentos (não fugir do assunto); que procede com lógica, conseqüente; abordagem de um assunto, os parágrafos relacionados.	
³ O nível de informatividade é o quanto o sumário se aproxima do texto original, se ele mostra a idéia principal do texto.	

No questionário foi verificada a precisão da sumarização de cada cenário sob dois critérios: o nível de informatividade e de qualidade. O nível de informatividade se refere a correlação do contexto do sumário em relação ao texto fonte. “Qualidade remete

¹ <http://www.grow.ic.ufal.br/>

ao fato de ser um texto e, assim, ter que apresentar as propriedades inerentes desse objeto lingüístico: minimamente, o sumário deve apresentar textura, isto é, deve ser coerente e coeso” (Rino e Pardo, 2006, p. 02).

O processo foi desenvolvido nas seguintes etapas: seleção do material a ser sumarizado (seja ela arquivo textual ou blog); sumarização manual dos materiais selecionados por parte dos autores do artigo; seleção da ferramenta de sumarização automática; sumarização automática dos materiais selecionados; elaboração do questionário de avaliação; aplicação do questionário avaliativo aos acadêmicos; análise quantitativa dos resultados; apresentação dos resultados obtidos e elaboração das conclusões.

Foi estabelecido um limite máximo para as sumarizações humanas: um terço do total de linhas do texto original. A ferramenta eleita para a sumarização automática é a Summarize Thingiee. Ela é online, implementada na linguagem php², gratuita e apresenta melhores resultados processando textos na língua inglesa. Além disso pode ser adicionada em qualquer site. Disponível através do link: <<http://www.tools4noobs.com/summarize/>>.

Dentre os parâmetros disponíveis na Summarize Thingiee, optamos por definir: o nº de linhas, o comprimento da frase: medido através do nº de caracteres, onde na maioria das vezes foi utilizado o padrão (50), além do comprimento da palavra em nº de caracteres mínimo (normalmente com o padrão: 4). Raramente, dependendo da necessidade, chegamos a usar a opção taxa de compressão (*Threshold*), que leva em consideração o nº de palavras relevantes na sentença. Mas como na maioria das vezes o nº de linhas foi especificado, o valor medido em porcentagem de compressão, foi desconsiderado. Ainda optamos por marcar a opção para visualizar as melhores palavras-chave. Como foi marcada a opção anterior, é liberada a opção de definir o nº destas palavras, neste caso definimos de acordo com o nº de palavras-chave da sumarização manual: entre 3 e 5, dependendo do tamanho do texto e o grau de relevância destas palavras.

5. Demonstração dos Resultados

A seguir abordaremos os resultados obtidos no experimento avaliativo da sumarização automática em comparação com a manual, em duas etapas: de repositórios de arquivos textuais e blogs.

Inicialmente observamos com a aplicação do questionário que os alunos da pesquisa estão na faixa etária de 18 a 21 anos, e 90% deles pertencem ao sexo masculino. Também foi questionado ao entrevistado qual das sumarizações apresentadas seria a automática, onde 80% deles conseguiu identificar corretamente. Os entrevistados, quando questionados a respeito da troca da sumarização manual pela automática, demonstraram um certo receio, preferindo ainda a manual com 75% dos votos para repositórios e 70% para os blogs.

Em relação às palavras-chave, constatou-se que as mesmas apresentam uma relação de proximidade ao documento original exatamente igual nas duas sumarizações (automática e manual), com 50%.

² <http://www.php.net>

a) Repositório de Arquivos Textuais

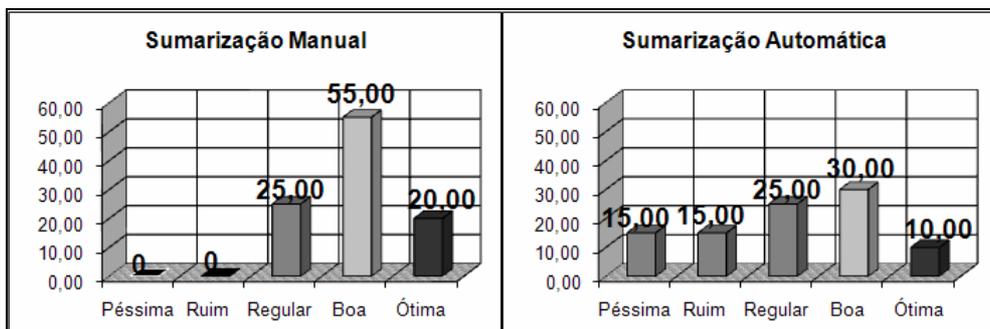


Figura 1- Como você definiria o nível de qualidade em relação à coerência das palavras da Sumarização?

Na Figura 1, pudemos constatar que na sumarização manual, mais da metade dos alunos considerou “Boa” o nível de qualidade da coerência das palavras. Enquanto, que na sumarização automática tivemos resultados balanceados, apesar do conceito “Boa” ter se sobressaído, como também ocorreu na Sumarização Manual.

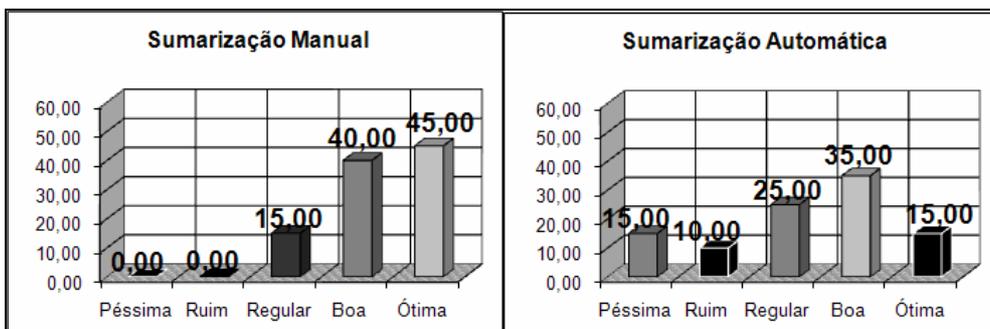


Figura 2- Como você definiria o nível de informatividade da Sumarização em relação ao texto original?

Já na Figura 2, averiguamos que o teor de exatidão em relação ao texto original ainda é inferior na sumarização automática, se comparada à manual. No entanto os resultados da sumarização automática ainda são satisfatórios, correspondendo à maioria dos votos.

b) Blog

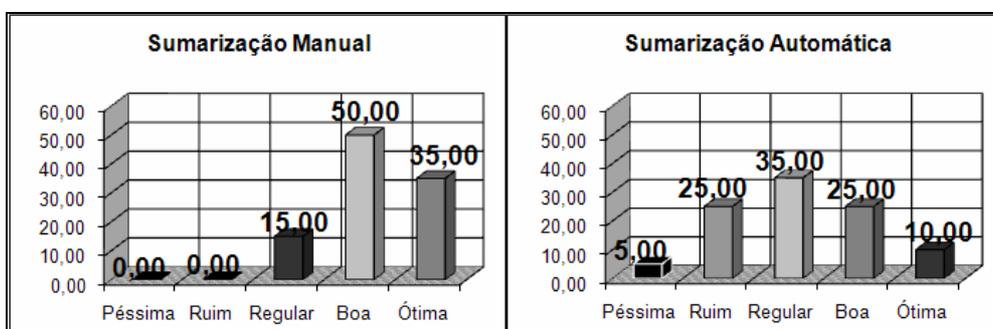


Figura 3 - Como você definiria o nível de qualidade em relação à coerência das palavras da Sumarização?

Aqui, o que se percebe na Figura 3, é que a coerência com o contexto do blog revela que a maioria dos votos a considerou mediana, através do conceito "Regular". Também podemos perceber que este último conceito, é igualado se somarmos os conceitos positivos (35%): "Ótima" e "Boa", mesmo que com uma pequena saliência, em relação à soma dos conceitos negativos (30%): "Ruim" e "Péssima". Aprovando desta forma, a sumarização automática.

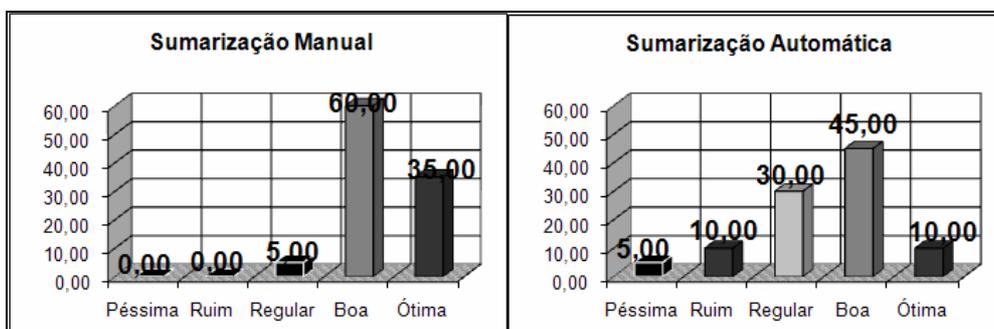


Figura 4 – Como você definiria o nível de informatividade da Sumarização em relação ao texto original?

A respeito do nível de informatividade, encontrada na sumarização das postagens do blog, a Figura 4 revela que a soma dos dois conceitos mais relevantes: "Ótima" e "Boa", correspondente a 55%, aprova a sumarização automática.

6. Conclusão

Perante os estudos realizados, os dados levantados e os resultados obtidos, concluímos que existem atualmente sumarizadores automáticos muito fiéis ao conteúdo, no entanto a produção desses não é tão próxima ao entendimento humano, causando dificuldade de compreensão e textualidade.

Ainda assim, temos um ganho de produtividade significativo ao utilizar um sumarizador automático se comparado às horas desperdiçadas com as sumarizações manuais. Focando no meio educacional, percebe-se a partir dos cenários aqui trabalhados: repositório de arquivos textuais e blogs, que os níveis de informatividade e qualidade em relação à coerência das palavras não sofreram perdas significativas nem consideráveis. Logo, temos uma alternativa viável a ser incorporada em Ambientes Educacionais na Web.

Apesar de a alternativa ter sido viável, é importante mencionar que enfrentamos a dificuldade da ferramenta utilizada trazer melhores resultados para sumários na língua inglesa. Espera-se então que se moldada ao português brasileiro os dados possivelmente trariam resultados mais satisfatórios.

Vale-se informar que devido a complexidade de manipular a ferramenta de sumarização automática em salas de bate papo (tratado no início deste artigo) deixaremos esse cenário para trabalhos futuros, assim como a implementação desse e do repositório de arquivos textuais e blogs.

Outro trabalho futuro consiste em agregar a sumarização automática a Ambientes Educacionais na Web, sistemas inteligentes otimizados para um melhor desempenho educacional operando de acordo com os cenários descritos neste artigo.

Referências

- Balage Filho, Pedro Paulo; Pardo, Thiago Alexandre Salgueiro; Nunes, Maria das Graças Volpe. (2007) “Experimentos com Sumarização Automática Extrativa de Textos Científicos”, In: I Workshop de Iniciação Científica e Tecnológica da Computação - ICMC - USP, São Carlos. Disponível em: <<http://www.icmc.usp.br/~tasparado/WICT2007-BalageEtAl.pdf>>. Acessado em: 01º de agosto de 2009.
- Bertagnolli, S. C. de; Silveira, Sidnei Renato da; Barcelos, Lauren Aparecida ; Nunes, E. (2007) “O Uso de Atividades Semipresenciais em Cursos Presenciais como Forma de Qualificação da Educação Superior: o caso do UniRitter”, In: RENOTE - Revista Novas Tecnologias na Educação, v. 5, nº. 1. Disponível em: <<http://www.cinted.ufrgs.br/ciclo9/artigos/11cSilvia.pdf>>. Acessado em: 7 de julho de 2009.
- Coracini, Sandra Regina. (2009) “O resumo como parâmetro de avaliação da compreensão leitora”, 171f. Dissertação de Mestrado, Departamento de Letras. Pontifícia Universidade Católica do Rio de Janeiro. Disponível em: <http://www2.dbd.puc-rio.br/pergamum/tesesabertas/0710561_09_pretextual.pdf>. Acessado em: 01º de agosto de 2009.
- Cristian, Radu. (2009) “Tools4noobs: Summarize Thingiee”. Disponível em: <<http://www.tools4noobs.com/summarize/>>. Acessado em: 06 de julho de 2009.
- Ferreira, Aurélio Buarque de Holanda. (2004) “Mini Aurélio: O Dicionário da Língua portuguesa”, 6ª Edição Revista e Atualizada, Curitiba: Posigraf, 896 p.
- Fuks, Hugo; Gerosa, Marco Aurélio; Pimentel, Mariano Gomes. (2003) “Projeto de Comunicação em Groupware: Desenvolvimento, Interface e Utilização”, In Congresso da Sociedade Brasileira de Computação, 23, Campinas, v. 2, p. 295-338. Disponível: <http://groupware.les.inf.puc-rio.br/groupware/publicacoes/JAI2003_cap7.pdf>. Acessado em: 23 de janeiro de 2008.
- Gonçalves, Patricia Nunes. (2008) “CorrefSum: Revisão de Coesão Referencial em Sumários Extrativos”, 136f. Dissertação de Mestrado. UNISINOS - Universidade do Vale do Rio dos Sinos, São Leopoldo. Disponível em: <http://btd.unisinos.br/tde_busca/arquivo.php?codArquivo=663>. Acessado em: 24 de julho de 2009.
- Hernandes, Carlos A. M.; Santana, Roberto A. S.; Falcão, Sérgio D. (2000) “Sobre o uso do chat como ferramenta auxiliar de ensino e aprendizagem no curso de Mestrado em Informática da Universidade Católica de Brasília”, Universidade Católica de Brasília, In: Revista Tecnologia da Informação, Ed. Universa, v. 2, n. 1, Brasília, p. 51-58. Disponível em: <<http://carlosmamede.org/Artigo%20sobre%20chat%20na%20UCB%20-%20publicado.pdf>>. Acesso em: 16 de julho de 2009.
- Oliveira, Bruno Vilela. (2008) “Uma Análise de Estratégias de Sumarização Automática”, 81f, Dissertação (Mestrado em Engenharia Civil) - Universidade Federal do Rio de Janeiro. Disponível em:

-
- <<http://www.dominiopublico.gov.br/download/texto/cp047453.pdf>>. Acessado em: 16 de julho de 2009.
- Mani, Inderjeet; Bloedorn, Eric; Gates, Barbara. (1998) “Using cohesion and coherence models for text summarization”, In: Symposium on Intelligent Text Summarization, 13, p. 69-76. Disponível em: <http://complingone.georgetown.edu/~linguist/papers/Spring98_Symp.pdf>. Acesso em: 26 de março de 2006. apud Marques Neto, Manoel Carvalho; Nascimento, Alisson B.; Gomes, Andréa Azevedo (2007) “SATSumm - Uma Ferramenta para Sumarização Automática de textos Jornalísticos”, In: ERBASE.
- Marques Neto, Manoel Carvalho; Nascimento, Alisson B.; Gomes, Andréa Azevedo (2007) “SATSumm - Uma Ferramenta para Sumarização Automática de textos Jornalísticos”, In: ERBASE - Escola Regional de Computação Bahia-Alagoas-Sergipe, 7, Vitória da Conquista, Bahia. Disponível em: <<http://manoelnetom.googlepages.com/artigoErbaseApresentado.pdf>>. Acessado em 21 de junho de 2009.
- Moresco, Silvia Ferreto da Silva; Behar, P. A. (2006) “BLOGS para a Aprendizagem de Física e Química”, In: RENOTE - Revista Novas Tecnologias na Educação, Porto Alegre, v. 4, n. 1. Disponível em: <http://www.cinted.ufrgs.br/renote/jul2006/artigosrenote/a44_21233.pdf>. Acessado em: 16 de julho de 2009.
- Rino, L. H. M. ; Pardo, Thiago Alexandre Salgueiro . (2006) “A Coleção TeMário e a Avaliação de Sumarização Automática”, Série de Relatórios do NILC. NILC-TR-06-04. São Carlos - SP: ICMC-USP. Disponível em: <<http://www.icmc.usp.br/~tasparado/NILCTR0604-RinoPardo.pdf>>. Acessado em 17 de julho de 2009.