

### III Workshop de Dissertações de Mestrado do PPgSI (2016)

#### FICHA DA PESQUISA

DADOS GERAIS					
Título do projeto de pesquisa	Reescrita sentencial baseada em traços de personalidade				
Orientando	Georges Basile Stavracas Neto				
Orientador(es)	Ivandr� Paraboni				
Semestre no curso, na data do workshop	<input type="checkbox"/> 2º semestre	<input checked="" type="checkbox"/> 3º semestre	<input type="checkbox"/> 4º semestre	<input type="checkbox"/> 5º semestre	<input type="checkbox"/> 6º semestre
Qualificação	<input type="checkbox"/> Qualificação já realizada em:		<input checked="" type="checkbox"/> Realização da qualificação planejada para: 31/10/2016		
Defesa	Prazo máximo para depósito: 05/02/2018		Realização da defesa planejada para: 30/11/2017		
Linha e Área de pesquisa	Gestão e desenvolvimento de Sistemas: <input type="checkbox"/> BD <input type="checkbox"/> Eng. de Software		Gestão de SI <input type="checkbox"/> IHC	Inteligência de Sistemas: <input checked="" type="checkbox"/> IA Padrões <input type="checkbox"/> Proc. Gráfico	
Área de aplicação	<input type="checkbox"/> Ambientes Corporativos <input type="checkbox"/> Bioinformática <input type="checkbox"/> Biometria <input type="checkbox"/> Economia	<input type="checkbox"/> Educação <input type="checkbox"/> Educação a Distância <input type="checkbox"/> Internet <input type="checkbox"/> Jogos	<input checked="" type="checkbox"/> Linguagem Natural <input checked="" type="checkbox"/> Linguística <input type="checkbox"/> Processos de Negócio <input type="checkbox"/> Química	<input type="checkbox"/> Redes Sociais <input type="checkbox"/> Robótica <input type="checkbox"/> Saúde <input type="checkbox"/> [outro – escrever]	
Publicações associadas ao projeto de mestrado	Não há publicações.				

#### DESCRIÇÃO DO PROJETO DE PESQUISA

Contextualização / motivação	Pesquisas recentes na psicologia fornecem o modelo Big-Five de personalidade, que é satisfatoriamente sólido e aplicável à computação, que pode usá-lo para modelar sistemas de língua natural baseados em personalidade.
Problema de pesquisa	O atual estado da arte possui poucas pesquisas que relacionam a geração de língua natural com personalidade do elocutor, e pouca ou nenhuma investigação no contexto brasileiro.
Objetivo geral da pesquisa	Investigar se é possível e como reescrever um texto em português alterando os traços de personalidade de quem o escreveu.
Trabalhos relacionados	<b>MAIRESSE, F.; Learning to adapt in dialog systems</b> mapeia os traços de personalidade e os parâmetros para gerar textos, mas só trabalha com dados-para-texto, e é específico para a língua inglesa. <b>MAIRESSE, F.; Trainable generation of Big-Five personality styles through data-driven parameter estimation</b> estende o trabalho anterior e possibilita usar mais valores como parâmetros para gerar textos, ainda com as mesmas limitações.
Justificativa e relevância	A expressão da personalidade em produções linguísticas (textos, fala, etc) é intrínseca à língua, mas não é um tópico extensamente estudado pela comunidade acadêmica como provado pela revisão sistemática realizada. Para aproximar sistemas que geram língua natural do desempenho humano, é necessário considerar a personalidade do emissor da mensagem.
Proposta para Solução	Será desenvolvido, com base nos dados já coletados, um modelo de reescrita sentencial que toma como parâmetros de entrada um texto qualquer, a personalidade de quem escreveu este texto e a personalidade-alvo para a qual o texto será reescrito.
Dados	Foram coletados dados na forma <i>Traços de personalidade</i> → <i>Exemplos de textos</i> para poder mapear quais traços de personalidade afetam quais parâmetros da geração de texto em português.
Validação	Os resultados serão avaliados intrinsecamente, comparando os textos gerados com a base de dados coletada. A depender do andamento da pesquisa, os dados também avaliados de maneira extrínseca por voluntários selecionados aleatoriamente.
Limitações	Os dados coletados tem poucos exemplos de personalidades extremas, e o conjunto é desbalanceado. A pesquisa só trata de reescrita de textos, e não trabalha com dados-para-texto.
Resultados esperados	Contribuições científicas: será gerado o modelo de reescrita de textos com base em traços de personalidade, que abrange a técnica para reescrever textos e os parâmetros que devem ser modificados. A base de dados coletada para a pesquisa também será liberada para a comunidade acadêmica no tempo oportuno. Contribuições tecnológicas: para ilustrar e validar o modelo, será desenvolvida a ferramenta “GNLP” e será

