

## IV Workshop de Dissertações de Mestrado do PPgSI (2017)

### FICHA DA PESQUISA<sup>i</sup>

DADOS GERAIS				
<b>Título do projeto de pesquisa</b>	Seleção de conteúdo referencial com base em traços de personalidade.			
<b>Orientando</b>	Danielle Sampaio Monteiro.			
<b>Orientador(es)</b>	Prof. Dr. Ivandré Paraboni.			
<b>Semestre no curso, na data do workshop</b>	<input type="checkbox"/> 2º semestre	<input type="checkbox"/> 3º semestre	<input checked="" type="checkbox"/> 4º semestre	<input type="checkbox"/> 5º semestre
<b>Qualificação</b>	<input checked="" type="checkbox"/> Qualificação já realizada em: 27/06/2017 <input type="checkbox"/> Realização da qualificação planejada para:			
<b>Defesa</b>	Prazo máximo para depósito: 29/08/2018		Realização da defesa planejada para: 20/02/2018	
<b>Linha e Área de pesquisa</b>	Gestão e desenvolvimento de Sistemas: <input type="checkbox"/> BD <input type="checkbox"/> Gestão de SI <input type="checkbox"/> Eng. de Software		Inteligência de Sistemas: <input checked="" type="checkbox"/> IA de Padrões <input type="checkbox"/> Proc. Gráfico	
<b>Área de aplicação</b>	<input type="checkbox"/> Ambientes Corporativos	<input type="checkbox"/> Educação	<input checked="" type="checkbox"/> Linguagem Natural	<input type="checkbox"/> Redes Sociais
	<input type="checkbox"/> Bioinformática	<input type="checkbox"/> Educação a Distância	<input type="checkbox"/> Linguística	<input type="checkbox"/> Robótica
	<input type="checkbox"/> Biometria	<input type="checkbox"/> Internet	<input type="checkbox"/> Processos de Negócio	<input type="checkbox"/> Saúde
	<input type="checkbox"/> Economia	<input type="checkbox"/> Jogos	<input type="checkbox"/> Química	<input type="checkbox"/> [outro – escrever]
<b>Publicações associadas ao projeto de mestrado</b>	Título do artigo: Personality-dependent referring expression generation. (Aceito) Autores: Ivandré Paraboni, Danielle Monteiro, Alex Lan. Título da conferência: TSD – Text, Speech and Dialogue			

### DESCRIÇÃO DO PROJETO DE PESQUISA

<b>Contextualização / motivação</b>	O trabalho de pesquisa está relacionado a um projeto de maior do professor Doutor Ivandré Paraboni área de Geração de Língua Natural com base em traços de personalidade.
<b>Problema de pesquisa</b>	O problema do trabalho está relacionado a área de Geração de Expressões de Referência gerar expressões referenciais semelhantes às produzidas por humanos, para isso intenta-se desenvolver um modelo de seleção de conteúdo para GER, baseado em traços de personalidade.
<b>Objetivo geral da pesquisa</b>	Desenvolver um modelo de seleção de conteúdo para geração de expressões de referência que leva em consideração traços de personalidade de um locutor alvo de interesse, e que produza descrições próximas das produzidas por humanos.
<b>Trabalhos relacionados</b>	Os trabalhos estudados dividiram-se em dois tipo: os algoritmos tradicionais com saídas fixas para uma dada entrada, como por exemplo o Algoritmo Incremental proposto por Dale e Reither, que serviram para modelar a tarefa de gerar expressões de referência; e os modelos que levam em consideração a variação humana para prever o comportamento referencial de cada indivíduo, como por exemplo os trabalhos de Dale e Viethan, e Ferreira, que utilizaram técnicas de aprendizagem de máquina para modelar suas propostas.
<b>Justificativa e relevância</b>	Algoritmos tradicionais de GER que produzem saída fixa não refletem a variação humana, já modelos que levam em conta variação individual fazem isso, mas exigem coleções de exemplos linguísticos produzidos por todos os indivíduos de interesse, tornando-se de difícil implementação e generalização. Como solução intermediária entre modelos de GER fixos e os baseados em locutores individuais, consideramos neste trabalho o uso de modelos que usam informações mais genéricas de grupos de pessoas com comportamento referencial similar.
<b>Proposta para Solução</b>	A proposta da presente pesquisa objetiva a definição, implementação e validação de um modelo para personalizar de forma genérica grupos de pessoas por meio da personalidade. Para auxiliar o desenvolvimento do modelo utilizaremos um cópulo com expressões de referência e dados de personalidade dos participantes.
<b>Dados</b>	O projeto contou com uma etapa de coleta de dados de expressões de referência produzidas e de dados de personalidade dos participantes por meio de um aplicativo de Facebook, o qual foi baseado em um inventário do modelo de personalidade <i>Cinco Grandes Fatores</i> (CGF)
<b>Validação</b>	A hipótese será testada comparando-se o modelo a ser proposto com o resultado da implementação do Algoritmo Incremental (algoritmo clássico para seleção de conteúdo) e aplicado ao corpus criado durante a pesquisa, b5-ref, por meio de métricas tradicionais da área, como os coeficientes de Dice e Acurácia.

<b>Limitações</b>	O trabalho tratará apenas da etapa de seleção de conteúdo, não abordando a tarefa de realização textual de um sistema GER, e as expressões geradas serão avaliadas somente na forma intrínseca utilizando a base de dados do <i>córpus de referência</i> , <i>b5-ref</i> , proposto pelo grupo de pesquisa local, a qual utiliza o modelo de personalidade CGF.
<b>Resultados esperados</b>	Contribuições científicas: Modelo viável de personalizar a seleção de conteúdo para gerar expressões de referência por meio do uso de personalidade do locutor. Contribuições tecnológicas: O <i>córpus</i> coletado, sendo o primeiro <i>córpus</i> de expressões de referência com os dados de personalidade dos participantes anotados.

<b>MÉTODO DE PESQUISA</b> [para cada item, marque todas as opções adequadas para seu projeto] (basear-se em: <a href="http://www.ufrgs.br/cursopgdr/downloadsSerie/derad005.pdf">http://www.ufrgs.br/cursopgdr/downloadsSerie/derad005.pdf</a> )			
<b>Gênero</b>	<input type="checkbox"/> Pesquisa teórica	<input checked="" type="checkbox"/> Pesquisa prática	<input type="checkbox"/> Pesquisa empírica <input type="checkbox"/> Pesquisa metodológica
<b>Natureza</b>	<input type="checkbox"/> Pesquisa básica/pura <input checked="" type="checkbox"/> Pesquisa aplicada		
<b>Objetivo</b>	<input type="checkbox"/> Pesquisa exploratória	<input type="checkbox"/> Pesquisa descritiva	<input checked="" type="checkbox"/> Pesquisa explicativa <input type="checkbox"/> [outro – escrever]
<b>Abordagem</b>	<input checked="" type="checkbox"/> Pesquisa quantitativa	<input type="checkbox"/> Pesquisa qualitativa	<input type="checkbox"/> Pesquisa mista (quali-quant)
<b>Procedimento(s) técnico(s)</b>	<input checked="" type="checkbox"/> Pesquisa experimental <input checked="" type="checkbox"/> Pesquisa bibliográfica <input type="checkbox"/> Pesquisa documental <input type="checkbox"/> Pesquisa <i>ex-post-facto</i> <input type="checkbox"/> Pesquisa de levantamento	<input checked="" type="checkbox"/> Pesquisa com <i>survey</i> <input type="checkbox"/> Estudo de caso <input type="checkbox"/> Pesquisa participante <input type="checkbox"/> Pesquisa-ação <input type="checkbox"/> Pesquisa etnográfica	<input type="checkbox"/> Pesquisa netnográfica <input type="checkbox"/> Teoria fundamentada em dados ( <i>grounded theory</i> ) <input type="checkbox"/> Ciência do projeto ( <i>Design science research</i> ) <input type="checkbox"/> [outro – escrever]
<b>Fonte(s) de dados</b>	<input type="checkbox"/> pesquisa de laboratório	<input checked="" type="checkbox"/> pesquisa de campo	<input checked="" type="checkbox"/> pesquisa bibliográfica <input type="checkbox"/> [outro – escrever]
<b>Técnica(s) / Instrumento(s) de coleta de dados</b>	<input type="checkbox"/> medição <input checked="" type="checkbox"/> questionário <input type="checkbox"/> entrevista <input type="checkbox"/> grupos focais	<input type="checkbox"/> formulário <input type="checkbox"/> <i>benchmark</i>	<input type="checkbox"/> observação (direta / participante) <input type="checkbox"/> diário de campo / notas de campo <input type="checkbox"/> análise documental (ou de artefatos) <input type="checkbox"/> [outro – escrever]
<b>Técnica(s) de análise de dados</b>	Análise quantitativa: <input checked="" type="checkbox"/> Estatística descritiva <input type="checkbox"/> Estatística inferencial <input type="checkbox"/> [outro – escrever]	Análise qualitativa: <input type="checkbox"/> Análise de conteúdo <input type="checkbox"/> Análise do discurso <input type="checkbox"/> [outro – escrever]	

<b>CRONOGRAMA</b> [altere conforme necessário]																																															
	2016												2017												2018																						
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12											
Revisão bibliográfica					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X																								
Coleta do <i>córpus</i>	X	X	X	X	X	X	X	X	X	X	X	X																																			
Anotação do <i>córpus</i>										X	X	X	X																																		
Desenvolvimento de Algoritmos Baseline													X																																		
Exame de qualificação																		X																													
Desenvolvimento do modelo proposto																			X	X	X	X	X																								
Validação do modelo																						X	X																								
Escrita da dissertação																		X	X	X	X	X	X	X	X																						
Depósito da dissertação																																															