IV Workshop de Dissertações de Mestrado do PPgSI (2017)

FICHA DA PESQUISA[i]

DADOS GERAIS					
Título do projeto de pesquisa	Ontologia voltada aos dados essenciais para estudos QSAR.				
Orientando	Maria Weslane de Sousa Almeida.				
Orientador(es)	Patricia Rufino Oliveira.				
Semestre no curso, na data do workshop	[ ] 2º semestre	[ ] 3º semestre	[ ] 4º semestre	[x] 5º semestre	[ ] 6º semestre
Qualificação	[x] Qualificação já realizada em: 06/12/2016 [ ] Realização da qualificação planejada para: dd/mm/aaaa				
Defesa	Prazo máximo para depósito: 05/02/2018 Realização da defesa planejada para: 22/01/2018				
Linha e Área de pesquisa	Gestão e desenvolvimento de Sistemas: [ ] BD[ ] Gestão de SI [ ] Eng. de Software [ ] IHC		Inteligência de Sistemas: [x] IA [ ] Rec. de Padrões [ ] Proc. Gráfico		
Área de aplicação	[ ] Ambientes Corporativos [ ] Bioinformática [ ] Biometria [ ] Economia	[ ] Educação [ ] Educação a Distância [ ] Internet [ ] Jogos	[ ] Linguagem Natural [ ] Linguística [ ] Processos de Negócio [ ] Química	[ ] Redes Sociais [ ] Robótica [ ] Saúde [x] Quimioinformática	
Publicações associadas ao projeto de mestrado	Submetida: Título do artigo: Applying transfer learning technique on QSAR models Autores: Rodolfo da Silva Simões, Maria Weslane de Sousa Almeida, Patrícia Rufino Oliveira e Clodoaldo Aparecido M. Lima Conferência: IEEE International Conference on BioInformatics and BioEngineering Em elaboração: Título do artigo: Chemical Data Problems When Applied in Machine Learning Techniques Autores (incluindo necessariamente o orientador): Maria Weslane de Sousa Almeida, Rodolfo da Silva Simões, Patrícia Rufino Oliveira, Káthia Maria Honório, Rafaela Molina de Ângelo e Sheila Cruz Araújo Conferência: IEEE International Conference on BioInformatics and BioEngineering				

DESCRIÇÃO DO PROJETO DE PESQUISA	
Contextualização / motivação	Estudos voltados à química medicinal necessitam de informações dos compostos referentes às propriedades químicas e aos ensaios biológicos que estão disponibilizados em diversas bases de dados.
Problema de pesquisa	Por os dados serem provenientes de diversas bases de dados, nem sempre é fácil determinar se apresentam correspondência entre eles, uma vez que, não existe um modelo capaz de descrevê-los adequadamente.
Objetivo geral da pesquisa	Descrever conceitos e relacionamentos de compostos químicos de acordo com sua atividade biológica, associando as informações de acordo com o método de obtenção (cálculo) dos descritores químicos importantes para estudos em QSAR (Quantitative Structure Analysis Relationships) para determinado alvo.
Trabalhos relacionados	Conceitos de química medicinal: MONTANARI, Carlos A. Química Medicinal. Métodos e fundamentos em planejamento de fármacos. São Paulo: Editora da USP, 2011. Formalização de ensaios biológicos: VEMPATI, Uma D. et al. Formalization, annotation and analysis of diverse drug and probe screening assay datasets using the BioAssay Ontology (BAO). PloS one, v. 7, n. 11, p. e49198, 2012. Formalização de dados químicos: HASTINGS, Janna et al. The chemical information ontology: provenance and disambiguation for chemical data on the biological semantic web. PLoS One, v. 6, n. 10, p. e25513, 2011.
Justificativa e relevância	Devido ao embasamento teórico suportado por uso de uma ontologia, os dados que estão dispostos em diversas bases de dados podem possuir a mesma anotação e formalização. Além de auxiliar pesquisadores que não possuem conhecimento teórico na química a compreender a relevância dos dados.

Proposta para Solução	(i) será proposta uma abordagem que descreva os dados químicos e biológicos dos compostos químicos; (ii) integração com ontologias já existentes.
Dados	Para o projeto de pesquisa foi criado um conjunto de dados de compostos químicos a fim de organizar os dados de acordo com informações biológicas, químicas e dos descritores QSAR.
Validação	Avaliar a estrutura da proposta de acordo com a documentação desenvolvida por meio de um frame, e também por especialistas da área fim.
Limitações	Os conjuntos de dados de Química Medicinal podem não possuir os descritores já calculados. Contudo a falta de dados não impede que os demais dados possam ser padronizados, o que servirá para formalizar as anotações referentes aos dados que poderão ser incluídos posteriormente.
Resultados esperados	Contribuições científicas: Espera-se que a integração e formalização dos dados químicos e biológicos suportado pela proposta auxiliem os estudos em QSAR no planejamento de medicamentos.
	Contribuições tecnológicas: Espera-se que a ontologia por meio de sua representação formal e da integração dos dados químicos e biológicos auxilie na aplicação de técnicas de aprendizagem de máquina e na criação de uma base de dados.

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

CRONOGRAMA [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
Estudo da literatura				x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x													
Revisão sistemática					x	x	x	x	x	x	x																									
Definição da proposta						x	x	x																												
Execução da proposta													x	x	x	x	x	x	x	x	x	x	X													
Coleta/obtenção de dados															x	x	x	x	x	x	x															
Análise dos dados																x	x	x	x	x	x															
Validação da proposta																	x	x	x	x	x	x	x													
Preparo da qualificação								x	x	x	x																									
Exame de qualificação												x																								
Escrita da dissertação													x	x	x	x	x	x	x	x	x	x	X													
Escrita de artigo																	x	x	x	x	x	x	x	x												
Depósito da dissertação																								x												

(Outros)																																				

i Esta ficha é uma adaptação da usada no "VIII Workshop de Teses e Dissertações em Sistemas de Informação (WTDSI 2015)" realizado como parte do "XI Simpósio Brasileiro de Sistemas de Informação (SBSI 2015)"