



<b>Evento</b>	Salão UFRGS 2013: SIC - XXV SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
<b>Ano</b>	2013
<b>Local</b>	Porto Alegre - RS
<b>Título</b>	Extração automática de correlação estratigráfica a partir de descrições de poços usando ontologias de domínio
<b>Autor</b>	LUAN FONSECA GARCIA
<b>Orientador</b>	MARA ABEL

Este trabalho se insere em um projeto que investiga abordagens automáticas de correlação estratigráfica a partir da descrição de poços de sondagem mineral. Estas descrições são realizadas utilizando uma ontologia de domínio já existente, que especifica conhecimento detalhado sobre litologias, texturas e estruturas para rochas sedimentares, ígneas e metamórficas.

Uma hipótese inicial deste trabalho é a de que é possível adaptar algoritmos clássicos de alinhamento de sequências de DNA, do domínio da bioinformática, para este problema específico, no domínio da Geologia. Para isso, é necessário desenvolver abordagens para determinar medidas de similaridade entre fácies sedimentares (elementos constituintes das descrições de poços). Tais medidas serão extraídas pelo uso de algoritmos de mineração de dados, aplicados sobre um conjunto de dados reais do domínio da Geologia. Estes dados, coletados pela CPRM (Companhia de Pesquisa de Recursos Minerais) em uma campanha para estudo das jazidas carboníferas do estado do RS, foram traduzidos do seu formato original para o padrão imposto pela ontologia, com o suporte de um sistema para descrição de testemunhos.

A metodologia adotada prevê (1) o estudo e teste de abordagens de data mining para determinar a similaridade entre descrições de fácies e (2) o estudo, adaptação e teste de algoritmos de alinhamento de DNA para o problema de correlação de poços. Até a data do envio deste resumo, foram realizados testes utilizando o algoritmo de clusterização EM (Expectation Maximization). Espera-se utilizar os clusters obtidos por esta abordagem para determinar a similaridade entre fácies sedimentares. A qualidade destes clusters será avaliada junto a geólogos. Nos próximos passos deste trabalho, estes resultados serão utilizados para adaptar os principais algoritmos de sequenciamento de DNA, tais como os algoritmos de Needleman-Wunsch e de Smith-Waterman