



Evento	Salão UFRGS 2020: SIC - XXXII SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
Ano	2020
Local	Virtual
Título	Busca Integrada de Dados Heterogêneos na Indústria do Petróleo: uma contribuição geocientífica
Autor	LUIZA GONÇALVES MAGGI
Orientador	CASSIANA ROBERTA LIZZONI MICHELIN

Busca Integrada de Dados Heterogêneos na Indústria do Petróleo: uma contribuição geocientífica

Autora: Luiza Gonçalves Maggi¹

Orientadora: Cassiana Roberta Lizzoni Michelin¹

¹ UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Ao longo das últimas décadas, a quantidade de dados geológicos adquiridos tem crescido exponencialmente, no entanto, quanto mais dados são gerados, maior é o esforço para organizá-los. Parte significativa das informações geológicas da indústria do petróleo está organizada de forma não estruturada, desta forma, uma solução de busca integrada, investigada nesse projeto, deve ser capaz de lidar com dados estruturados (tabelas de banco de dados), e dados não estruturados (texto, imagem, vídeo e áudio). Este trabalho tem como objetivo principal auxiliar, através do domínio de geologia, a procura de soluções para a busca integrada de dados heterogêneos que utilizam a semântica da informação. Para atingir este objetivo, este trabalho realizou um levantamento do domínio pelos boletins de geociências da Petrobras, banco de teses e dissertações e sites de busca onde se verificou as principais imagens empregadas na indústria do petróleo. Foi proposta uma ontologia das imagens, seguindo uma classificação com os principais tipos de representações. Verificou-se que as imagens podem ser agrupadas em quatro grupos principais (imagem, diagrama, gráfico e fotografia) e 46 subgrupos (mapas, perfis, seções, dentre outros) de acordo com o objeto a ser representado. Logo após, foram realizados testes online de classificação de imagens com o intuito de se obter um grande banco de dados de imagens geológicas com base na ontologia proposta. Esses testes foram realizados por professores, alunos e profissionais que atuam no domínio geocientífico. Foram classificadas, por 66 usuários, três mil imagens. Foi utilizado o algoritmo Machine Learning, que conseguiu classificar uma imagem do domínio da geociências com 80% de acurácia dentre 24 classes da ontologia proposta. Desta forma, este trabalho ressalta a importância da contribuição geológica, em conjunto com as outras áreas, para a futura construção de uma plataforma de busca integrada dos dados geocientíficos heterogêneos que será utilizada pela empresa Petrobras.