

9º CONGRESSO BRASILEIRO DE PESQUISA E DESENVOLVIMENTO EM PETRÓLEO E GÁS



9º CONGRESSO
Brasileiro de P&D em
PETRÓLEO E GÁS

Maceió, AL
de 09 a 11 de novembro
2017

Realização:



ABPG
ASSOCIAÇÃO BRASILEIRA DE P&D EM
PETRÓLEO E GÁS



TÍTULO DO TRABALHO:

DATA SCIENCE FOR OIL: ESTADO DA ARTE E PROPOSTA DE FERRAMENTA COMPUTACIONAL BASEADA EM CIÊNCIA DE DADOS PARA INDÚSTRIA DE PETRÓLEO E GÁS

AUTORES:

Gustavo Coelho Lopes, Eduardo Setton Sampaio da Silveira, Josué Domingos da Silva Neto, Heitor Soares Ramos Filho

INSTITUIÇÃO:

Engenharia de Petróleo, Centro de Tecnologia, Universidade Federal de Alagoas,
josuedsneto@gmail.com, 2Engenharia Civil, Centro de Tecnologia, Universidade Federal de Alagoas

Este Trabalho foi preparado para apresentação no 8º Congresso Brasileiro de Pesquisa e Desenvolvimento em Petróleo e Gás - 9º PDPETRO, realizado pela Associação Brasileira de P&D em Petróleo e Gás - ABPG, no período de 09 a 11 de novembro de 2017, em Maceió/AL. Esse Trabalho foi selecionado pelo Comitê Científico do evento para apresentação, seguindo as informações contidas no documento submetido pelo(s) autor(es). O conteúdo do Trabalho, como apresentado, não foi revisado pela ABPG. Os organizadores não irão traduzir ou corrigir os textos recebidos. O material conforme, apresentado, não necessariamente reflete as opiniões da Associação Brasileira de P&D em Petróleo e Gás. O(s) autor(es) tem conhecimento e aprovação de que este Trabalho seja publicado nos Anais do 9º PDPETRO.

DATA SCIENCE FOR OIL: ESTADO DA ARTE E PROPOSTA DE FERRAMENTA COMPUTACIONAL BASEADA EM CIÊNCIA DE DADOS PARA INDÚSTRIA DE PETRÓLEO E GÁS

Abstract

The stages of exploitation, development and production of oil and gas produce a great volume of data (Spath, 2014). Data science is the interdisciplinary field responsible for extracting knowledge from data using principles and techniques of computing, engineering, statistics, economics, among other disciplines, to create innovative solutions through automatic data analysis. This data, that can be structured, non-structured or even semi-structured, is not yet broadly used, specifically because of the challenge that is deciding which data is relevant.

Big Data applications are in its early developments in the oil and gas industry, however, according to Spath (2014), with the increase of field and well monitoring, specially due to the pressure being made by regulatory agencies over the world, data analysis is the next great technology innovation in the oil industry. Mehta (2016) cites that data analysis is one of the great priorities for oil and gas companies in 2018.

In Brazil, data regarding the oil industry is under the responsibility of the Nacional Oil Agency (ANP) and can be found in reports, tables and bulletins, that way it lacks a centralized database with intuitive access. Beyond ANP, there are reports that can be found in initiatives brought by organizations such as OPEC, BP, EIA, the initiatives are JODI, Knoema, Dataviva or even in commercial softwares as the RystadEnergy. One of the problems found in the industry and research centers is related to the time-consuming task of finding, obtaining and aggregating the data for further analysis. In an environment where decision making has to be near instantaneous, time is a crucial matter.

This work presents the analysis of the development of a computational database for oil and gas – Data Science for Oil, which aims at obtaining, treating and analyzing the data that is broadly available in the internet and the databases of regulatory agencies and oil and gas companies to provide a complete and integrated database for oil and natural gas in Brazil.

Introdução

Estamos na quarta revolução industrial, Wang et al. (2016), Bloem et al. (2014), e os dados estão aqui para ficar. Só resta às indústrias tradicionais aceitarem essa nova revolução ou arriscarem serem consideradas atrasadas.

Na era da informação, o modo mais confiável de se tomar decisões se tornou o Big Data (Lohr, 2012), que se trata do tratamento de dados que possuem certas características em comum, sendo as principais, o Volume, a Variedade e a Velocidade na qual essas informações se tornam disponíveis. Nessa nova forma de tratar os dados faz-se uma análise histórica e em tempo real para auxílio da tomada de decisão, o período de tomar decisões baseadas em instinto está com seus dias contados.

O volume de dados tratados na indústria do óleo e gás sempre foi historicamente grande, dados contidos em data centers de empresas modernas de óleo e gás podem chegar à grandeza de 20 petabytes (Beckwith, 2011), principalmente agora com o advento da Internet das Coisas e sensores com informação em tempo real. Esses dados sempre foram usados para auxiliar a tomada de decisão, porém em uma menor escala. Não há o uso sistemático e contínuo dos dados para tomar decisões mais fundadas (Perrons, 2015).