



UNIVERSIDADE FEDERAL DE PERNAMBUCO  
GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO  
CENTRO DE INFORMÁTICA  
2016.1

Análise exploratória de dados: um estudo diagnóstico sobre a satisfação de clientes  
de um banco através de Ontologias e Web Semântica

**PROPOSTA DE TRABALHO DE GRADUAÇÃO**

Aluna: Déborah Mesquita (dhsm@cin.ufpe.br)

Orientadora: Bernadette Farias Lóscio (bfl@cin.ufpe.br)

Recife, 27 de Abril de 2016

## 1. Contextualização

Hoje em dia costuma-se dizer que "dados são o novo petróleo". Com isso a profissão de data scientist ganhou bastante destaque, principalmente depois do famoso artigo *Data Scientist: The Sexiest Job of the 21st Century* (Thomas H. Davenport e D.J. Patil), publicado na Harvard Business Review em 2012 [1]. Assim como o petróleo, os dados precisam ser refinados, e é justamente nesta etapa de "refino" onde se observa o papel do cientista de dados.

Os *insights* e respostas proporcionados pela análise de dados servem como enorme fonte de vantagem competitiva, fazendo com que as empresas que efetuem decisões a partir dos dados apresentem um desempenho muito melhor do que as que não fazem uso destas informações.

O trabalho dos *data scientists* é extrair conhecimento dos dados, uma tarefa semelhante a tarefa da Web Semântica. Fazer com que a semântica dos dados possa ser interpretada por computadores abre um novo caminho para o desenvolvimento de soluções envolvendo a análise de dados. Com o uso de ontologias (modelagem formal da estrutura de um sistema) e da Web Semântica é possível que se represente o conhecimento advindo de diversas fontes de dados.

A Web Semântica também representa uma forma de diminuir o abismo entre os dados e as respostas que os estrategistas das empresas precisam para tomar decisões. Utilizando-se ontologias e processamento de linguagem natural é possível que o processo de análise dos dados seja feito de forma automática, tornando as informações acessíveis para os funcionários e facilitando uma possível publicação destas informações na Intranet da empresa.

## 2. Objetivos

Devido às redes sociais, a satisfação dos clientes é um fator cada vez mais importante para empresas. A mudança do paradigma de Marketing de transação (cujo principal objetivo são as transações de compra e venda) para o paradigma de Marketing de relacionamentos (onde o foco é gerar valor para os clientes) exige que as empresas ofereçam um atendimento personalizado para cada cliente, identificando pontos positivos e negativos da relação entre cliente e empresa.

Neste trabalho de graduação o objetivo é encarar o desafio que as empresas possuem em transformar os dados em informações relevantes, com o objetivo de oferecer um melhor serviço a seus clientes. Utilizando conceitos de Semântica e Ontologias pretende-se analisar os dados de clientes do Banco Santander (base de dados obtida através da plataforma de competições Kaggle).

A análise envolverá o entendimento do domínio (satisfação de clientes e serviços bancários), o entendimento dos dados, definição das perguntas que deverão ser respondidas com os dados e modelagem uma ontologia para representar o conhecimento. Também será estudada e desenvolvida

uma forma de se visualizar os dados, com a criação de um website onde os funcionários da empresa poderão fazer consultas. A maior contribuição do trabalho é a criação de um *framework* que possibilite a transformação de dados sobre clientes e transações bancárias em uma plataforma de consulta e visualização de informações, que represente através de uma ontologia as características e os níveis de satisfação de cada cliente. Com isso se espera que os funcionários do banco sejam capazes de oferecer um atendimento mais personalizado, baseado nas necessidades de cada cliente.

### 3. Cronograma

<b>Atividade</b>	<b>Março</b>	<b>Abril</b>	<b>Maió</b>	<b>Junho</b>	<b>Julho</b>
Formulação da proposta e definição do problema					
Revisão da Literatura					
Entendimento do domínio, criação das perguntas e modelagem da ontologia					
Análise e categorização dos dados					
Criação da plataforma para visualização dos dados					
Elaboração do relatório final					
Preparação para defesa					

### 4. Possíveis Avaliadores

Os possíveis avaliadores deste trabalho de graduação serão:

- Ana Carolina Salgado
- Bernadette Farias Lóscio

## 5. Referências

[1] DAVENPORT, Thomas H; PATIL, D.J . Data Scientist: The Sexiest Job of the 21st Century. Disponível em: <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century/>. Acesso em: 23 de Março de 2016.

[2] LORANG, Noah. Practical skills that practical data scientists need. Disponível em: <https://m.signalvnoise.com/practical-skills-that-practical-data-scientists-needda71e6b93f95#.k4sp3enu1>. Acesso em: 23 de Março de 2016.

[3] GRANVILLE, Vincent. Developing Analytic Talent: Becoming a Data Scientist. John Wiley & Sons, 2014.

[4] WOODS, Dan. How Semantics Can Make Data Analysis Work Like A Google Search. Disponível em: <http://www.forbes.com/sites/danwoods/2013/11/26/how-semantics-can-make-dataanalysis-like-a-google-search/#2c57d70d28ed>. Acesso em: 28 de Março de 2016.

[5] DAVIES, John; GROBELNIK, Marko; MLADENIC, Dunja. Semantic Knowledge Management. Springer, 2009.

## 6. Assinaturas

---

Déborah Mesquita

**Orientanda**

---

Bernadette Farias Lóscio

**Orientadora**