



ANTONIO MENEGHETTI FACULDADE - AMF
CURSO DE SISTEMAS DE INFORMAÇÃO

ROGER SILVEIRA

USO DE BIG DATA PARA ANÁLISE DE DADOS EM REDES SOCIAIS

RESTINGA SECA/RS

2016

ROGER SILVEIRA

USO DE BIG DATA PARA ANÁLISE DE DADOS EM REDES SOCIAIS

Trabalho de Conclusão do Curso de Sistemas de Informação para obtenção do título em Bacharel de Sistemas de Informação.
Orientador: Profº. Fábio Sarturi Prass.

RESTINGA SECA/RS

2016

AGRADECIMENTOS

Primeiramente a Deus, por esta oportunidade.

A instituição Antonio Meneghetti Faculdade, seu corpo docente, direção e administração que proporcionaram um ambiente extremamente agradável e de grande aprendizado.

Ao meu orientador, o profº **Fábio Sarturi Prass**, pelo empenho e tempo dedicado à elaboração deste trabalho.

A todos os professores do curso de sistemas de informação, que foram importantes na minha trajetória acadêmica e no desenvolvimento deste trabalho. Um agradecimento especial ao profº **José Luiz R. Filho** que foi muito importante na escolha do tema para a elaboração deste trabalho e desde o começo se propôs a me ajudar e tirar dúvidas.

A **FOIL** e sua metodologia com a qual recebi ensinamentos que são e serão muito importantes tanto na minha vida pessoal quanto na minha carreira profissional, ensinamentos estes que me motivam a fazer e ser cada vez melhor, me tornando um líder em cada decisão e oportunidade.

Ao meu pai **Gedernei Silveira**, minhas irmãs **Francine Silveira** e **Fabrine Silveira** e também a minha namorada **Stefane Unfer** que, com muito carinho e apoio, não mediram esforços para que eu chegasse até esta etapa de minha vida.

A minha mãe **Nibia I. J. Silveira** que sempre me ensinou a fazer o meu melhor e ir em busca dos meus sonhos e objetivos.

E a todos que de alguma forma fizeram parte da minha formação, o meu muito obrigado!

“Tente uma, duas, três vezes e se possível tente a quarta, a quinta e quantas vezes for necessário. Só não desista nas primeiras tentativas, a persistência é amiga da conquista. Se você quer chegar à onde a maioria não chega, faça o que a maioria não faz.”

(BILL GATES)

RESUMO

SILVEIRA, Roger. **O uso de Big Data para análise de dados em redes sociais**. 2016. Trabalho de conclusão de curso. Faculdade Antonio Meneghetti. Curso de Sistemas de Informação, Recanto Maestro-Restinga Seca/RS, 2016.

O objetivo deste trabalho é apresentar os conceitos do termo Big Data, realizando na prática a coleta e análise dos dados para obter informações em tempo real. A coleta de dados é realizada nas redes sociais Facebook e Twitter, utilizando APIs disponibilizadas pelas mesmas. No Facebook são coletadas informações considerando publicações, comentários e curtidas em postagens. Já no Twitter são coletadas informações que contenham termos também relacionados, ambas coletas com o objetivo de demonstrar como Big Data pode auxiliar na tomada de decisões, obtendo essas informações em tempo real, tornando-se assim um diferencial no mercado.

Palavras-chave: Big Data, coleta de dados, grande volume de dados, tomada de decisão.

ABSTRACT

SILVEIRA, Roger. **O uso de Big Data para análise de dados em redes sociais**. 2015. Trabalho de conclusão de curso. Faculdade Antonio Meneghetti. Curso de Sistemas de Informação, Recanto Maestro-Restinga Seca/RS, 2015.

The aim of this paper is to present the concepts of the term Big Data, realized the practice of collecting and analyzing data to get information in real time. Data collection is done on social networks Facebook and Twitter, using APIs provided by them. Facebook is collected information given posts, comments and tanned on posts. Already on Twitter is collected information containing terms also related, both collections in order to demonstrate how big data can assist in making decisions, obtaining this information in real time, thus becoming a differentiator in the market.

Keywords: Big Data, data collection, large amounts of data, decision making.

LISTA DE ABREVEATURAS

APIs - Application Programming Interface

HTML - HyperText Markup Language

JSON - JavaScript Object Notation

LINQ - Language Integrated Query

NoSQL - Not Only Structured Query Language

RFID - Radio Frequency Identification

SQL - Structured Query Language

URL - Uniform Resource Locator

LISTA DE ILUSTRAÇÕES

Figura 1 - O que acontece em um minuto na internet.	16
Figura 2 - Representação de relacionamento entre tabelas.....	20
Figura 3 - Tela de criação do aplicativo no Facebook.....	24
Figura 4 - Painel de controle da aplicação no Facebook.	25
Figura 5 - Obtendo o identificador da página da Fiat Automóveis Brasil.....	26
Figura 6 - Retorno de informações sobre a página da página Fiat Automóveis Brasil.	27
Figura 7 - Visualizando <i>feeds</i> coletados.	28
Figura 8 - Identificando o atributo <i>paging</i>	28
Figura 9 - Coletando curtidas e comentários em uma publicação.	29
Figura 10 - Página para criação de aplicativos no Twitter.	30
Figura 11 - Página para informar dados do aplicativo.....	30
Figura 12 - Página para obter as chaves e access token de autenticação no Twitter.....	31
Figura 13 - Retorno da consulta no Twitter.....	32
Figura 14 - Informações sobre o usuário da publicação.....	33
Figura 15 - Publicações negativas referente ao Fiat Mobi.	36
Figura 16 - Publicações positivas referente ao Fiat Mobi.	36
Figura 17 - Gráfico de publicações por dia da semana.....	38
Figura 18 - Gráfico de publicações por hora.	38

SUMÁRIO

LISTA DE ABREVEATURAS	7
LISTA DE ILUSTRAÇÕES	8
1 INTRODUÇÃO	10
1.1 OBJETIVOS	11
1.1.1 Objetivo Geral	11
1.1.2 Objetivo Específico	11
1.2 JUSTIFICATIVA	12
2 ABORDAGEM TEÓRICA	13
2.1 O QUE É BIG DATA?	13
2.2 BENEFÍCIOS E RISCOS DO BIG DATA	14
2.3 DADOS - ORIGENS E FORMATOS	15
2.4 COLETA DE DADOS EM REDES SOCIAIS	18
2.5 BANCO DE DADOS RELACIONAL <i>VERSUS</i> BANCO DE DADOS NoSQL	19
3 METODOLOGIA	22
4 ESTUDO DE CASO	23
4.1 COLETA DE DADOS NO FACEBOOK	23
4.2 COLETA DE DADOS NO TWITTER	29
4.3 APLICANDO BIG DATA NA TOMA DE DECISÕES	33
5 CONCLUSÃO	40
REFERÊNCIAS	41

1 INTRODUÇÃO

Hoje em dia grande parte da população tem acesso a internet através de um celular, *tablet* ou computador. Só no Brasil a pessoas passam em média nove horas por dia por dia conectadas, de acordo com uma pesquisa realizada pela agência internacional We Are Social (DAMASCENO, 2015).

Tudo isso acaba resultando em um grande acúmulo de dados, gerados através de diversas fontes e armazenado de diferentes formas. Grande parte destes elementos correspondem aos dados não estruturados, dados estes que não podem ser organizados em linhas e colunas, ou seja, dados que até então só podiam ser compreendidos por seres humanos, como: vídeos, fotos, postagens em redes sociais, e-mails, registros de ligações entre outros.

As empresas também contam com um grande volume de dados, segundo Taurion (2013) as organizações trabalham com um volume de dados superior a uma dezena de terabytes e em alguns casos até mesmo superior a alguns petabytes, porém isso não significa que quanto mais dados uma empresa tiver mais conhecimento ela vá possuir. Para adquirir esse conhecimento elas necessitam de ferramentas com capacidade de coletar e analisar estes dados em tempo real, para que essa informação obtida se transforme em conhecimento capaz que auxiliar no dia-a-dia.

Converter estes dados em informação relevante para conseguir um melhor conhecimento, não é mais possível utilizando as ferramentas tradicionais de TI. Com as ferramentas adequadas para tratamento dessas essas informações, situadas em um grande volume de dados, essas informações podem se tornar em conhecimento que pode resultar oportunidades de negócio, se traduzindo em potenciais benéficos econômicos para as empresas.

De acordo com Alecrim (2013), utilizando as ferramentas adequadas, Big Data oferece a possibilidade de converter essa informação em conhecimento e assim gerar novas

oportunidades de negócio e dar visibilidade do mercado, ajudando a conquistar novos clientes e também na tomada decisões em tempo real.

O conceito de Big Data pode ser definido como sendo um conjunto de dados extremamente amplos e complexos, que por este motivo, necessitam de ferramentas especialmente preparadas para lidar com grandes volumes, de uma forma que toda e qualquer informação encontrada seja analisada e aproveitada em tempo hábil. Big Data pode ser compreendido como a análise de grandes quantidades de dados para a geração de resultados importantes que, em volumes menores, dificilmente seriam alcançados (ALECRIM, 2013).

Segundo Uehara (2015), com o Big Data uma organização torna-se apta a conseguir respostas que humanamente levariam séculos para serem obtidas, exatamente pela capacidade das ferramentas de processar altos volumes de dados e extrair padrões inesperados.

1.1 OBJETIVOS

1.1.1 Objetivo Geral

Utilizar Big Data afim de encontrar resultados específicos, que possam auxiliar na tomada de decisões, acerca de uma permissa sendo executada em redes sociais.

1.1.2 Objetivo Específico

1. Conceituar sobre Big Data (características e aplicabilidade).
2. Apresentar ferramentas que auxiliam na coleta e análise de dados.
3. Realizar na prática a coleta e análise de um grande volume de dados.
4. Mostra como os resultados obtidos podem auxiliar na tomada de decisões.

1.2 JUSTIFICATIVA

Com o rápido crescimento da tecnologia, e o aumento exponencial da geração de dados, se tornada cada vez mais essencial observar e compreender as ferramentas, com as quais se possam analisar e trabalhar com esse grande volume de dados, o que pode se tornar um diferencial na tomada de decisões.

2 ABORDAGEM TEÓRICA

O Big Data marca o início de uma importante transformação. Como várias outras tecnologias que se tornaram o diferencial para empresas competitivas no mercado, o Big Data desafia o modo em que o mundo é visto até hoje, alterando a forma de como as decisões são tomadas e até mesmo a compreensão da realidade (SCHÖNBERGER, 2013).

De acordo com Davenport e Kin (2014), o Big Data e a analítica nele baseada prometem mudar praticamente todos os setores de atividade e todas as funções de negócios. Qualquer negócio ou profissional que estiver inserido neste assunto conquistará vantagens competitivas importantes.

2.1 O QUE É BIG DATA?

O conceito de Big Data é focado no grande armazenamento de dados que acontece com enorme velocidade, a medida com que são gerados, esse termo é baseado no conceito 5V (Valor, Veracidade, Variedade, Volume e Velocidade) e tem sido tratado por muitos, como a solução de eventuais situações problemáticas onde a informação em tempo real poderá contribuir (ADAMI, 2011).

De acordo com Alecrim (2013), Big Data pode ser definido como sendo o conjunto de dados extremamente amplos e que, por este motivo, necessitam de ferramentas especialmente preparadas para lidar com grandes volumes de dados, de uma forma que toda a informação disponível possa ser encontrada, analisada e aproveitada em tempo hábil.

"Big Data se refere a trabalhos em grande escala que não podem ser feitos em escala menor, para extrair novas ideias e criar novas formas de valor de maneira que alterem os

mercados, organizações e até relações entre cidadãos e governos" (SCHÖNBERGER, 2013).

Segundo Santana (2013), o grande desafio do Big Data é administrar um grande volume de dados e obter informações em um menor tempo de requisição. Com o grande volume de dados, fazer com que a aplicação cresça à medida que é necessário é uma ótima estratégia, assim, uma escalabilidade vertical (em que se aumenta o poder do hardware, como aumento de memória e de processamento de uma única máquina) ou horizontal (em que se aumenta a quantidade de máquinas) deve ser analisada.

Para Schönberger (2013), o Big Data se trata de uma nova fonte de valor econômico e inovação, representando mudanças na forma de como a informação é analisada, uma vez que com as ferramentas adequadas, se torna possível processar e analisar uma quantidade de dados muito maior que as ferramentas tradicionais suportam, com isso se torna possível enxergar detalhes antes nunca vistos por estar limitado a uma quantidade de dados menores.

2.2 BENEFÍCIOS E RISCOS DO BIG DATA

De acordo com Schönberger (2013), ao mesmo tempo que Big Data pode gerar grandes benefícios como disponibilizar informações preciosas em tempo real e proporcionar resultados mais amplos por poder contar com uma base de dados muito maior do que as empresas poderiam obter, essa ferramenta pode também oferecer alguns perigos.

Segundo Schönberger (2013), em dois mil e nove, quando um novo vírus de gripe, conhecido como H1N1, se espalhou rapidamente preocupando os órgãos de saúde pública de todo o mundo, temendo que esse novo vírus pudesse repetir casos como o da gripe espanhola, que em 1918 infectou meio bilhão de pessoas e matou dezenas de milhões.

Para tentar conter uma nova epidemia, os CDCs (Centros de Controle e Prevenção de Doenças) nos Estados Unidos pediram para que cada médico relatasse os novos casos de contaminação deste novo vírus para elaborar planos assim elaborar um plano para desacelerar a propagação deste vírus, já que até então não havia uma vacina disponível contra ele. Mesmo assim, esse plano desenhado contava com informações já defasadas, pelo fato de que o intervalo entre as pessoas começarem a sentir os sintomas e conseguirem realizar uma

consulta médica poderia levar dias, dias estes que se juntavam ao tempo de transferência das informações para as organizações centras (SHÖNBERGER, 2013).

Esse é um dos casos onde Big Data mostra um pouco do seu poder e seus benefícios. De acordo com Shönberger (2013), anteriormente ao vírus H1N1 ser conhecido a Google já havia desenvolvido um mecanismo que poderia prever a disseminação da gripe de inverno nos Estados Unidos, não apenas nacionalmente, mas também por regiões específicas e até mesmo estados. Como a Google recebe bilhões de pesquisas diárias e as armazena em suas bases de dados, a empresa pode realizar previsões ao analisar estas pesquisas e comparar com os dados do CDC sobre a disseminação da gripe entre 2003 e 2008, com a intenção de identificar pessoas contaminadas através de suas pesquisas na internet.

Assim, quando em 2009 a crise do vírus H1N1 ocorreu, o sistema da Google provou ser um indicador mais útil e pontual que as estatísticas do governo, munindo as autoridades de saúde pública com valiosas informações em tempo real (SHÖNBERGER, 2013).

De acordo com Shönberger (2013) boa parte dos dados gerados, inclui também, informações pessoais. Com isso, essa nova era da tecnologia e da internet pode colocar em risco a privacidade das pessoas, onde a cada aplicativo utilizado ou rede social acessada, as informações do usuário são liberadas para os fornecedores destas tecnologias. Com isso, pode ser possível descobrir os gostos de cada pessoa, locais frequentados e até mesmo onde a pessoa está naquele momento. Estes dados podem não parecer informações explícitas, porém com processos de análise, podem facilmente dizer a quem se referem ou facilitar a dedução de detalhes íntimos, se tornando possível até mesmo montar um perfil de cada pessoa.

De acordo com Shönberger (2013), big data pode se tornar um risco se utilizado para obter essas informações para uso inadequado ou até mesmo para prever ações de uma pessoa para julgá-las ou puni-la antes mesmo que elas hajam.

Para Shönberger (2013), big data é útil para entender o presente, os riscos futuros e para auxiliar na hora de decidir as ações adequadas a serem tomadas, suas previsões podem ajudar pacientes, médicos, empresários que precisam de informação em tempo real para uma tomada de decisão, porém, big data nada diz sobre as casualidades e por isso pode ser perigoso utilizar essa tecnologia na hora de julgar alguma casualidade e determinar a culpa de alguém.

2.3 DADOS - ORIGENS E FORMATOS

De acordo com a Cisco (2013), são geradas diariamente enormes quantidades de dados por *smartphones*, sensores, câmaras de vídeo, contadores inteligentes e outros dispositivos com ligação à internet, que se juntam à já significativa quantidade de informação gerada pelos meios tradicionais. Meios estes que podem ser representados por exemplo: por sistemas comerciais que acumulam uma grande quantidade de informações sobre seus clientes, fornecedores, operações, produtos, vendas, entre outras diversas.

Segundo Schönberge (2013), a Google processa mais de vinte e quatro petabytes por dia, valor este que poderia representar pelo menos quinze vezes ao armazenamento de dez bilhões de fotos. O Facebook, empresa com pouco mais de dez anos de existência, recebe por volta de dez milhões de fotos a cada hora, além de registrar cerca de três bilhões de dados referentes a curtidas e comentários de seus usuários. Dados estes que são capazes de criar uma trilha digital ajudando a descobrir as preferências de cada usuário.

A Figura 1 mostra uma representação feita pela Cisco (2013), de o que acontece na internet de um minuto, trazendo mais para perto da realidade o grande acúmulo de dados gerados na internet.

Figura 1 - O que acontece em um minuto na internet.



Fonte: Cisco (2013).

Todos esses dados podem ter diversas origens, entre elas: web e redes sociais, máquina a máquina, biometria, dados de transações, dados gerados por pessoas entre outros:

- Os dados obtidos da web e redes sociais são dados gerador por páginas web e redes sociais como o *Facebook*, *Twitter*, *Instagram*, *Youtube*, entre outras. Estes dados são representados por exemplo: fluxos de clicks, *tweets*, *feeds*, *postings* e conteúdos diversos da internet.

- Dados máquina a máquina, são gerados pelos objetos e dispositivos conectados a entre si ou a nuvem, entre os dispositivos empregados para capturar estes tipos de dados estão: os marcadores RFID¹, os sensores, dispositivos GPS entre outros.
- Dados gerados por biometria, a informação biométrica se refere a identificação automática de uma pessoa, baseada em suas características anatômicas ou traços pessoais, os dados anatômicos são criados a partir das características físicas de uma pessoa incluindo: impressão digital, scanner de retina, reconhecimento facial, informação genética e reconhecimento de voz.
- Dados de transações: essa categoria compreende os dados que são gerados em transações comuns como: compras com cartões de crédito, dados para o faturamento de serviços, registros detalhados de ligações, dados de reclamações nas empresas de seguro.
- Dados gerados pelas pessoas: as pessoas geram uma grande quantidade de dados como: ligações telefônicas, notas de voz, correios e documentos eletrônicos entre muitos outros.

Estes dados podem ser classificados em dados estruturados, dados sem-estruturados e dados não estruturados:

- Dados estruturados: são dados organizados em blocos semânticos, dados de um mesmo grupo possuem características semelhantes, ou seja, que possuem atributos. Esse tipo de dados tem uma lógica em uma estrutura identificadas, podem ser armazenados em banco de dados relacionais e podem ser agrupados em filas e colunas.
- Dados semiestruturados: são uma forma de dados estruturados que podem ser entendidos por máquinas, tem um formato e fluxo lógico, porém não são amigáveis para o usuário, como exemplo: códigos HTML², logs de sites Web, código XML.
- Dados não estruturados: dados que não podem ser guardados em um formato de filas e colunas como: arquivos de vídeos, arquivos de áudios, dados de sequências de clicks entre outros, esse tipo de dado representa a grande maioria dos dados em geral.

¹ (Radio Frequency Identification – Identificação por Radiofrequência) Tecnologias que utilizam a frequência de rádio para captura de dados.

² Linguagem de marcação utilizada na construção de páginas na Web.

2.4 COLETA DE DADOS EM REDES SOCIAIS

De acordo com Damasceno (2015), com a facilidade de troca de mensagens com amigos e familiares e o acesso rápido as informações proporcionadas pelos usuários, as redes sociais estão no topo em relação ao uso de internet nas maiores economias mundiais.

De acordo com Damasceno (2015), com base em previsões é possível estimar que a internet possa chegar ao acesso de mais de 50% da população mundial em 2016, sendo possível prever que as redes sociais cheguem ao número de pelo menos 2,1 bilhões de pessoas com perfis ativos.

As redes sociais continuam a crescer em todo o mundo e de acordo com um estudo realizado por uma agência de marketing social, a We Are Social, estima-se que a média de uso em redes sociais chegue a duas horas e vinte e cinco minutos por dia (DAMASCENO, 2015).

O Facebook e o Twitter são as redes sociais mais utilizadas. O Facebook lidera este requisito, com um número superior a um bilhão e meio de usuários ativos no ano de 2015. Já o Twitter, rede social que possibilita aos usuários a troca de atualizações pessoais através de textos de até 140 caracteres, conhecidos como *tweets*, mantém uma posição em as dez mais utilizadas (COMSCHOOL, 2015).

Com isso, as redes sociais se tornam um instrumento muito importante para adquirir informações relevantes sobre seus usuários, informações essas que pode ajudar empresas a deixar o seu produto o mais próximo possível do gosto de seu público alvo. Com simples informações como curtidas e comentários em postagens ou páginas no Facebook, pode ajudar também na hora de decidir em que local abrir uma nova filial e definir se o produto lançado está agradando seus usuários, entre várias outras informações que podem se tornar o diferencial no mercado competitivo.

Segundo França (2014), as principais redes sociais online disponibilizam interfaces ou serviços para a captura de seus dados, normalmente existem duas formas de coleta de dados, uma delas consiste em determinar termos e coletar as informações que façam referência ao mesmo, como por exemplo: buscar todas as postagens, publicações que contenham o termo “Brasil”, a segunda forma de coleta seria a captura das informações a média com que as mesmas vão surgindo, essa é uma maneira indicada para empresas que desejam monitorar a satisfação, de seus clientes, com seus produtos.

O Twitter é uma rede social online que possui duas APIs (*Application Programming Interface*) diferentes para a captura dos seus dados: REST API e *Streaming API*, para a utilização de ambas APIs é necessário inicialmente que o usuário tenha uma conta no Twitter. Estas APIs trabalham com o padrão de arquivo JSON (*JavaScript Object Notation*), onde todos os dados são recebidos nesse formato (FRANÇA, 2014).

O Facebook também disponibiliza algumas APIs de consumo, uma delas se chama *Graph API*. Para utilizar a *Graph API* é necessário criar uma aplicação que deverá ser vinculada a uma conta no Facebook (FRANÇA, 2014).

2.5 BANCO DE DADOS RELACIONAL VERSUS BANCO DE DADOS NoSQL

Os Bancos de Dados Relacionais foram desenvolvidos para prover acesso facilitado aos dados, possibilitando que os usuários utilizassem uma grande variedade de abordagens no tratamento das informações (FOWLER, 2013). Segundo Lima (2013), dentre os bancos de dados relacionais mais conhecidos estão o banco de dados Oracle, My SQL e SQL Server, ambos utilizam a linguagem que também é padrão para os demais bancos de dados relacionais que é a *Structured Query Language*, ou SQL, como é mais conhecida.

Segundo a IBM Corporation, o modelo relacional é composto de entidades e relacionamentos, onde uma entidade pode ser uma tabela física ou uma projeção lógica de diversas tabelas, as entidades por sua vez possuem colunas, cada uma identificada por um nome e um tipo.

De acordo com Leite (2010), no modelo relacional as informações são distribuídas em colunas e tabelas e onde cada linha se refere a um registro diferente como por exemplo: um sistema que guarde as informações de seus empregados, terá uma tabela que poderá conter informações como o nome e salário, cada um destes seria um campo nesta tabela, cada funcionário cadastrado seria uma linha desta tabela ou seja um novo registro. Esta tabela poderia também guardar registros que façam referência a uma outra tabela, para assim ter o relacionamento entre as mesmas.

A Figura 2 representa o relacionamento entre tabelas, onde a tabela “Empregado” guarda as informações dos empregados de uma determinada empresa, contendo informações como: código, nome e salário dos empregados, além do código de departamento que faz

referência a tabela “Departamento”, assim informando para qual departamento aquele empregado pertence.

Figura 2 - Representação de relacionamento entre tabelas.

Exemplo de um BD Relacional

Empregado

NumEmp	NomeEmp	Salario	Dept
032	J Silva	380	21
074	M Reis	400	25
089	C Melo	520	28
092	R Silva	480	25
112	R Pinto	390	21
121	V Simão	905	28
130	J Neves	640	28

Departamento

NumDept	NomeDept	Ramal
21	Pessoal	142
25	Financeiro	143
28	Técnico	144

Fonte: NETO (2008).

Segundo Brito (2011), o Modelo Relacional tem sido amplamente utilizado em praticamente todos os tipos de sistemas de bancos de dados nas últimas décadas. Porém, o crescimento cada vez mais rápido do volume de dados, têm propiciado que modelos alternativos de banco de dados passem a ser utilizados. Motivados principalmente pela questão da escalabilidade do sistema, uma nova geração de bancos de dados, conhecidos como NoSQL, começa a ganhar força e espaço tanto na academia quanto no mercado.

De acordo com Diana (2010) o mundo nunca havia lidado com volumes de dados tão grandes, em parte, graças a Web 2.0 passaram a ser encontrados limites nas técnicas e ferramentas de gerenciamento de dados disponíveis naquele momento, criando assim a necessidade de uma ferramenta capaz de suportar esse grande volume. Foi assim que o banco de dados NoSQL surgiu, pelo fato dos bancos de dados relacionais terem dificuldade em lidar com estes dados.

Para Fowler (2013) as grandes organizações estão percebendo a importância de armazenar a maior quantidade de dados possível e processá-los mais rapidamente. Um banco

de dados NoSQL além de ter a capacidade de atender estes requisitos, além de, em alguns casos se tornar mais barato que um banco de dados relacional, uma vez que a execução destes dados relacionais é projetada para a execução em uma única máquina, mas geralmente é mais econômico executar grandes quantidades de dados em clusters com máquinas menores e mais baratas.

O banco de dados NoSQL tem por objetivo de atender aos requisitos de gerenciamento de grandes volumes de dados, sejam eles semiestruturados ou não estruturados, que necessitam de alta disponibilidade e escalabilidade. Segundo Lóscio (2011), os bancos de dados NoSQL tem sido amplamente adotado em empresas como Facebook, Amazon e Google com a finalidade de atender às suas demandas de escalabilidade, alta disponibilidade e dados não estruturados.

No caso dos bancos de dados NoSQL, toda a informação é armazenada de forma agrupada no mesmo registro, ou seja, a informação fica toda em um mesmo registro ao invés de ter essa informação desmembrada no relacionamento de várias outras tabelas (FOUWLER, 2013).

O banco de dados NoSQL se torna também mais produtivo no desenvolvimento de aplicativos, pelo simples fato de que muito trabalho no desenvolvimento de aplicações é gasto no mapeamento de dados entre as estruturas de dados na memória e um banco de dados relacional. Um banco de dados NoSQL pode fornecer um modelo de dados que se adapte às necessidades de aplicativos, simplificando assim, essa interação e resultando em menos código a ser escrito e depurado, melhorado (FOUWLER, 2013).

3 METODOLOGIA

Este trabalho consiste abordar os conceitos de Big Data, além de realizar na prática a análise e coleta de dados mostrando como as informações que antes só podiam ser obtidas através de uma análise humana (levando um tempo muito maior de análise) e que com as ferramentas adequadas auxiliam empresas a obterem informações em tempo real, que possam ajudar na tomada de decisões.

Neste trabalho foi realizada a coleta de dados em duas redes sociais, Facebook e Twitter. A coleta no Facebook foi realizada baseada na página da Fiat no Brasil, coletando informações como postagens, curtidas e comentários relacionados ao novo carro da marca, lançado em abril deste ano, o Fiat Mobi. Já a coleta de dados no Twitter é realizada por postagens (*tweets*) que contenham termos também relacionados a este carro.

As coletas de dados foram realizadas utilizando APIs disponibilizadas pelas próprias redes sociais. Para coleta de dados no Facebook utilizou-se a *Graph API*, já para a coleta de dados no Twitter foi utilizada a *REST API* e *Streaming API*.

Após realizar a coleta de dados, também foi realizada uma análise nas informações obtida foi realizada, identificando como as mesmas podem auxiliar na tomada de decisões.

4 ESTUDO DE CASO

Para Siqueira (2010) é muito comum os clientes utilizarem as redes sociais para expressarem a sua opinião sobre determinado produto, eles criticam, elogiam, indicam e até mesmo realizam análises sobre este. Esse boca a boca em grandes proporções influencia diretamente as vendas do produto, podendo ser benéfico ou não.

O monitoramento em redes sociais fornece respostas para compreender e avaliar ações tomadas por empresas, como estas ações estão sendo conduzidas e recebidas por seus clientes. Essas respostas funcionam como forma de um feedback em grande escala e em tempo real, onde saber o que está sendo falado sobre sua marca, produtos e serviços é extremamente fundamental para o melhor posicionamento frente as mídias sociais.

Com isso, para este trabalho foram realizadas coletas afim de obter informações que possam dar alguns indícios sobre a satisfação dos consumidores frente a esse produto, obtendo também em tempo real a informação sobre algum comentário negativo, permitindo assim que a empresa possa, em menos tempo, executar alguma ação para prevenir que este comentário possa de alguma forma prejudicar a visão dos consumidores frente ao produto.

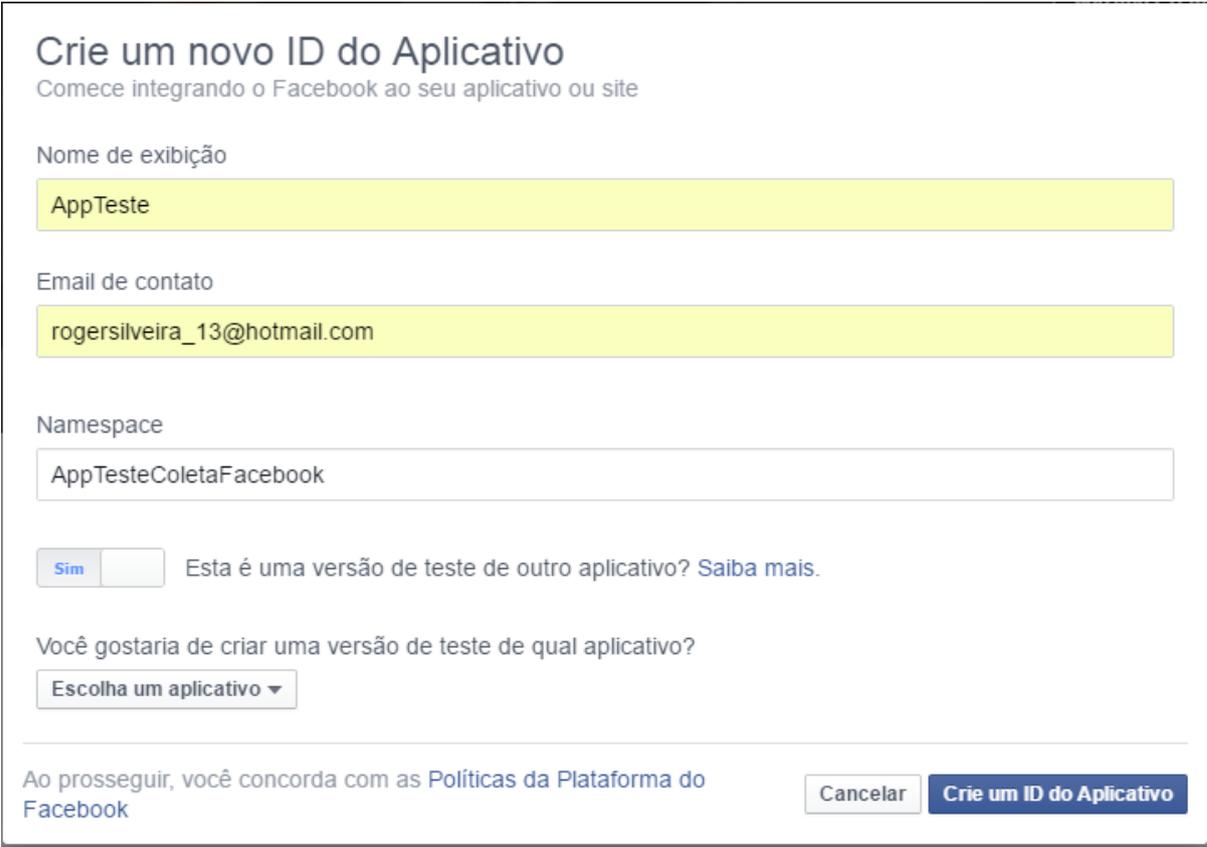
4.1 COLETA DE DADOS NO FACEBOOK

O Facebook disponibiliza um conjunto de rotinas para realizar a coleta de dados na rede social, essas rotinas são disponibilizadas através da Graph API, que trabalha com o retorno de dados no formato JSON. O JSON é um formato de troca de dados e informações entre sistemas muito utilizado por sua velocidade na execução e no transporte destes dados, além de sua simples leitura e seu tamanho reduzido.

Para começar a consumir a Graph API, deve-se inicialmente criar uma aplicação dentro do próprio Facebook acessando a página de desenvolvedor. A Figura 3 demonstra a tela de

criação do aplicativo, onde o usuário informa alguns dados como um nome de exibição para a aplicação, o e-mail que será utilizado para comunicados sobre o aplicativo e um identificador exclusivo, esse último sendo opcional.

Figura 3 - Tela de criação do aplicativo no Facebook.



The screenshot shows the Facebook 'Create a new App ID' form. The title is 'Crie um novo ID do Aplicativo' with the subtitle 'Comece integrando o Facebook ao seu aplicativo ou site'. The form contains three input fields: 'Nome de exibição' with the value 'AppTeste', 'Email de contato' with the value 'rogersilveira_13@hotmail.com', and 'Namespace' with the value 'AppTesteColetaFacebook'. Below these fields is a checkbox labeled 'Sim' with the text 'Esta é uma versão de teste de outro aplicativo? Saiba mais.' and a question 'Você gostaria de criar uma versão de teste de qual aplicativo?' with a dropdown menu showing 'Escolha um aplicativo'. At the bottom, there is a disclaimer 'Ao prosseguir, você concorda com as Políticas da Plataforma do Facebook' and two buttons: 'Cancelar' and 'Crie um ID do Aplicativo'.

Fonte: Facebook (2016)

Após a criação da aplicação o usuário tem acesso a uma chave de autenticação, conhecida como *access token*, essa chave é necessária para ler, modificar ou gravar dados no Facebook através desta API. Para ter acesso a essa chave o usuário deve executar um método passando informações como o identificador da aplicação criada e a uma chave secreta, ambas informações são geradas pelo Facebook após a criação da aplicação e podem ser obtidas acessando o painel de controle da aplicação assim como mostra a Figura 4. Para visualizar a chave secreta do aplicativo basta o usuário clicar no botão mostrar, disponibilizado junto ao campo e informar a senha da sua conta no Facebook.

Figura 4 - Painel de controle da aplicação no Facebook.



Fonte: Facebook (2016)

Para realizar a coleta dos dados neste trabalho foi desenvolvida uma aplicação na plataforma .Net, utilizando como linguagem de programação C#, para realizar a comunicação entre a aplicação e a API foi utilizada a classe *WebClient*, do próprio framework da plataforma.

A Listagem 1 mostra a execução do método para obter o *access token*, onde na linha 1 está sendo instanciado um objeto da classe *WebClient*, para assim ter acesso a função *DownloadString*, que é responsável por executar a chamada método e retornar à informação para a aplicação, assim como mostra a linha 5, onde é executado o método passando a URL declarada na linha 4 e atribuído o valor de retorno para a variável “*accessToken*”, que é utilizada no restante do sistema para executar os demais métodos.

A linha 4 mostra a declaração de uma variável que recebe a URL responsável por obter a informação do *access token*, passando como parâmetros o “*id*” da aplicação e a chave secreta.

Listagem 1 – Obtendo o *access token*.

1. `client = new WebClient();`
2. `string appId = "238348076215393";`
3. `string chaveSecreta = "p3vvvhf5942wa204d2832g586e66a845";`
4. `string oauthUrl = string.Format("https://graph.facebook.com/oauth/access_token?type=client_cred&client_id={0}&client_secret={1}", appId, chaveSecreta);`
5. `accessToken = client.DownloadString(oauthUrl).Split('=')[1];`

Após obter o *access token* o usuário já pode começar a realizar a coleta de dados. Para iniciar a coleta de informações em uma página deve-se inicialmente identificar de qual página

deseja coletar as informações. Para identifica-la deve ser passado o identificador da mesma, o qual pode ser obtido acessando o perfil da página no próprio Facebook, como pode ser visualizado na Figura 5.

Figura 5 - Obtendo o identificador da página da Fiat Automóveis Brasil.



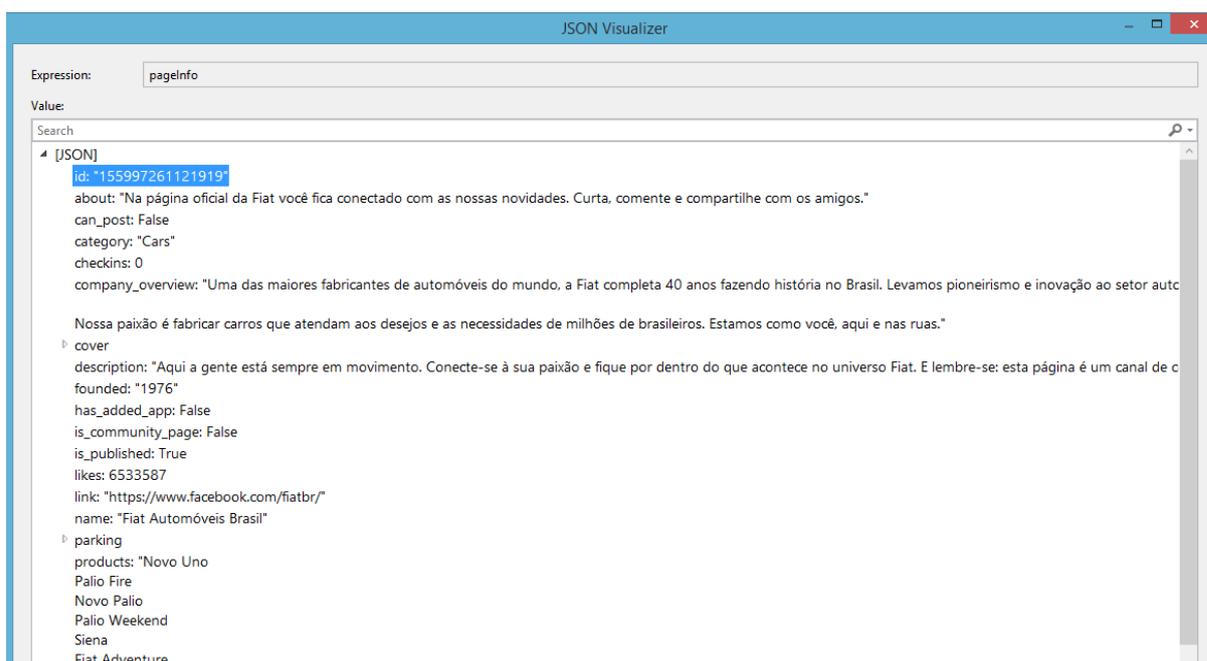
Fonte: Facebook (2016)

A Listagem 2 mostra o código utilizado para coletar informações referente a página, onde na linha 1 é declarada uma variável que recebe o identificador da página e na linha 2 é executado o método para coletar essas informações que são retornadas no formato JSON. Esse método retornar informações como o “id” da página, uma descrição da mesma, o número de usuários que curtiram a página, os produtos, entre outras informações assim como pode ser visto na Figura 6.

Listagem 2 – Obtendo informações sobre a página da página Fiat Automóveis Brasil.

1. `string indentPage = "fiatbr";`
2. `string pageInfo = client.DownloadString(string.Format("https://graph.facebook.com/{0}?access_token={1}", indentPage, accessToken));`

Figura 6 - Retorno de informações sobre a página da página Fiat Automóveis Brasil.



A Graph API permite também realizar a coleta de informações mais abrangentes como o *feed* e publicações da página, além das curtidas e comentários nestas publicações, permite também coletar informações como fotos, vídeos e até mesmo os usuário que curtiram a página ou alguma publicação da mesma, além de outras várias informações disponíveis.

Na Listagem 3 observa-se a execução do método para realizar a coleta de *feeds* da página e o retorno das informações, esse método retorna todas as publicações visíveis no *feed* da mesma, também no formato JSON, assim como pode ser visto na Figura 7. Porém neste caso as informações são retornadas de forma paginada, retornando de vinte e cinco em vinte e cinco e para obter as próximas informações deve-se executar a URL disponibilizada através do atributo *paging* junto ao retorno na execução do método, como pode ser visto na Figura 8.

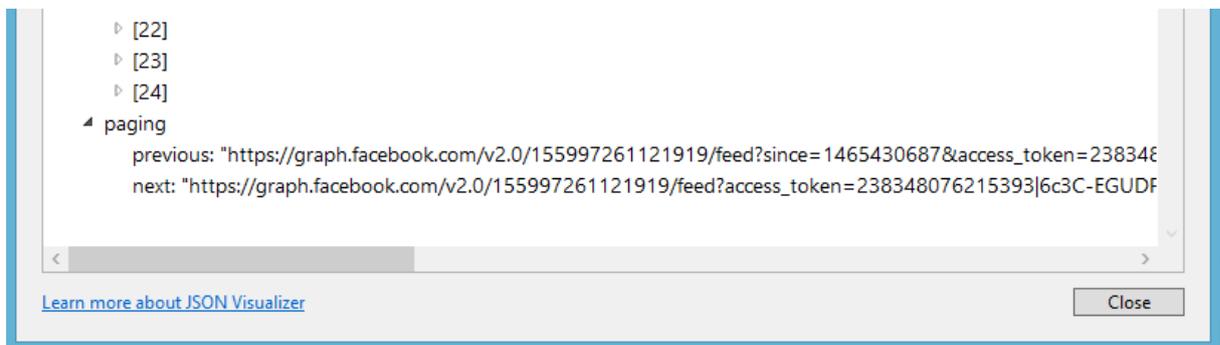
Listagem 3 – Coleta de *feed* da página.

1. `string url = string.Format("https://graph.facebook.com/{0}/feed?access_token={1} ", indentPage, accessToken);`
2. `string pagefeed = client.DownloadString(url);`

Figura 7 - Visualizando *feeds* coletados.

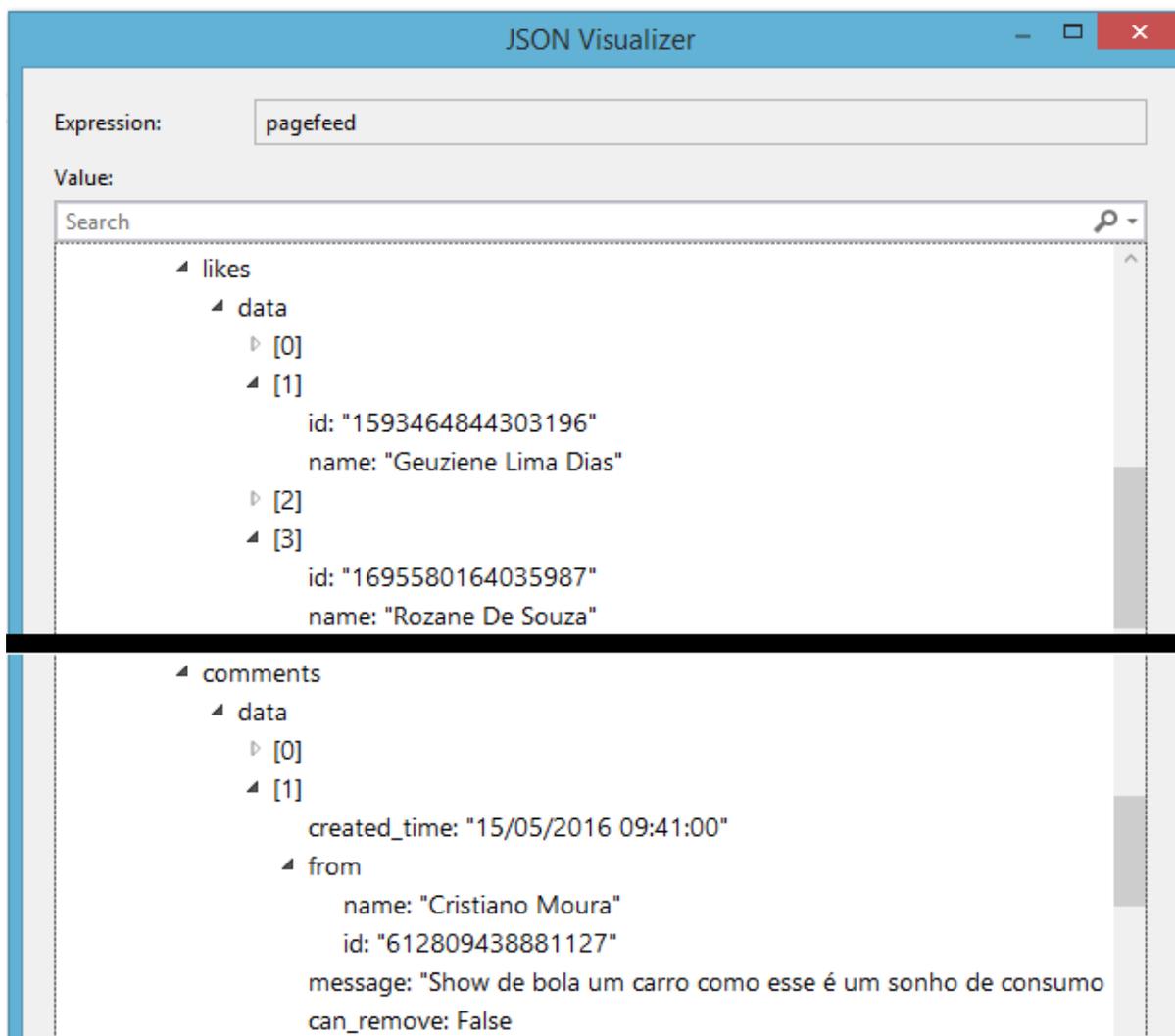


Figura 8 - Identificando o atributo *paging*.



Neste mesmo objeto de retorno é possível ter acesso a informações referente a curtidas e comentário dessa mesma publicação. Na Figura 9 observa-se o objeto de retorno com estas informações sobre curtidas e comentários, onde para as curtidas é possível obter informações como o nome e o "id" do usuário que curtiu a respectiva publicação e para os comentários, além destas mesmas informações é possível também obter o conteúdo do comentário publicado.

Figura 9 - Coletando curtidas e comentários em uma publicação.



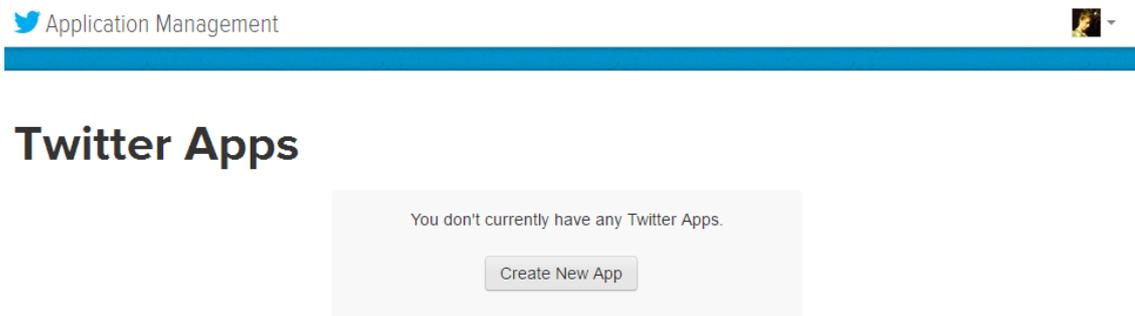
4.2 COLETA DE DADOS NO TWITTER

Para realizar a coleta de dados no Twitter foi utilizada a biblioteca LINQ to Twitter. De acordo com Mayo (2015), o LINQ to Twitter é um provedor LINQ *open source* para o Twitter (serviço de *micro-blogging*), que utiliza a sintaxe LINQ (*Language Integrated Query*) como padrão para realizar consultas na rede social, a interação entre a biblioteca e o Twitter é feita através da API do próprio Twitter.

De acordo com Macoratti (2012), o LINQ é uma extensão do .NET Framework que permite realizar consultas (Queries) diretamente na base de dados ou em objetos, documentos XML, estrutura de dados e coleção de objetos entre outros.

Para começar a coletar informações no Twitter o usuário precisa de uma autenticação, neste caso a autenticação também é feita através da criação de um aplicativo, que deve ser criado através da página de aplicativos do Twitter, assim como mostram as Figuras 10 e 11, onde a Figura 10 mostra a página inicial para a criação de aplicativos e a Figura 11 mostra a página para informar os dados referente ao aplicativo que está sendo criado.

Figura 10 - Página para criação de aplicativos no Twitter.



Fonte: Twitter (2016).

Figura 11 - Página para informar dados do aplicativo.

Create an application

Application Details

Name *

Your application name. This is used to attribute the source of a tweet and in user-facing authorization screens. 32 characters max.

Description *

Your application description, which will be shown in user-facing authorization screens. Between 10 and 200 characters max.

Website *

Your application's publicly accessible home page, where users can go to download, make use of, or find out more information about your application. This fully-qualified URL is used in the source attribution for tweets created by your application and will be shown in user-facing authorization screens.
(If you don't have a URL yet, just put a placeholder here but remember to change it later.)

Fonte: Twitter (2016).

Após a criação do aplicativo o usuário tem acesso a algumas chaves (*Consumer Key*, *Consumer Secret*, *Access Token* e *Access Token Secret*) que são necessárias para autenticar-se e começar a realizar a coleta de dados, o acesso a estas chaves pode ser obtido pela página “*Keys and Access Tokens*” assim como mostra a Figura 12.

Figura 12 - Página para obter as chaves e access token de autenticação no Twitter.

Aplicativo Coleta Teste Twitter

[Details](#)
[Settings](#)
[Keys and Access Tokens](#)
[Permissions](#)

Application Settings

Keep the "Consumer Secret" a secret. This key should never be human-readable in your application.

Consumer Key (API Key)	RApP756KiVfwBUoOHDphOFV0
Consumer Secret (API Secret)	T8qchO0POU6Prj7ZN1fKI54YVugvshXxl46Tc0ie4cmYJ9kXHG
Access Level	Read and write (modify app permissions)
Owner	silveiraroger
Owner ID	140972677

Fonte: Twitter (2016).

A Listagem 4 mostra o código utilizado para realizar a coleta de dados no Twitter, onde da linha 1 até a linha 12 está sendo declarado o objeto “authorizer” com as informações de autenticação do usuário, passando as chaves obtidas após a criação do aplicativo.

Na linha 13 está sendo declarado um objeto do tipo “TwitterContext”, onde está sendo passado por parâmetro o objeto “authorizer” que contém as informações de autenticação para a interação com o Twitter, o objeto “twitterContext” é utilizado para realizar a consulta no Twitter, onde é realizada uma consulta LINQ passando como condições o tipo de pesquisa e o termo que para filtrar os *twitts* que são relacionados a determinado assunto, no caso deste trabalho está consultando *twitts* que contém o termo “Fiat Mobi”, para assim obter as publicações relacionadas a este automóvel.

Listagem 4 – Coleta de dados no Twitter.

```

1. var authorizer =
2.     new SingleUserAuthorizer
3.     {
4.         CredentialStore =
5.             new SingleUserInMemoryCredentialStore
6.             {

```

```

7.         ConsumerKey = "RApP756KiVfwBUoOHDphOFFVo",
8.         ConsumerSecret = "T8qchO0POU6Prj7ZN1fKI54YVugvshXxl46Tc0ie4cmyJ",
9.         AccessToken = "14097ewq13fmQK8RdfewqeF8nNtdlcVf",
10.        AccessTokenSecret = "OSdn3usd8K32v1eWmHeY10f432gJbAL1jS"
11.    }
12. };
13. var twitterContext = new TwitterContext(authorizer);
14. var tweets = (from search in twitterCtx.Search
15.              where search.Type == SearchType.Search
16.              && search.Query == "\"Fiat Mobi\"")
17.              select search.Statuses
18.    ).ToList();

```

Nas Figuras 13 e 14 observa-se o retorno das informações consultadas através do código descrito a cima, obedecendo os parâmetros informados. Esse método retorna informações como conteúdo da publicação e informações relacionadas ao usuário que realizou a publicações, dentre estas informações estão: o nome do usuário, o local de onde foi publicado e até mesmo o idioma da publicação.

Figura 13 - Retorno da consulta no Twitter.

Name	Value
Scopes	Count = 0
ScreenName	null
SinceID	0
Source	"Twitter V
StatusID	736371747222630400
Text	"Fiat Mobi Muito Bom!"
TrimUser	false
Truncated	false
TweetIDs	null
Type	Home
User	{LinqToTwitter.User}
UserID	0
Users	Count = 0
WithheldCopyright	false
WithheldInCountries	Count = 0
WithheldScope	null

Figura 14 - Informações sobre o usuário da publicação.

Expression:
`(new System.Collections.Generic.Mscorlib_CollectionDebugView<LinqToTwitter.Status>(teste)).Items[0].User`

Value:

Name	Value
 FavoritesCount	0
 FollowersCount	413
 Following	false
 FollowRequestSent	false
 FriendsCount	172
 GeoEnabled	false
 ImageSize	Normal
 IncludeEntities	false
 IsTranslator	false
 LangResponse	"pt"
 ListedCount	2
 Location	"Restinga Seca"
 Name	"Roger Silveira"
 Notifications	false

4.3 APLICANDO BIG DATA NA TOMA DE DECISÕES

Simple ações na internet como uma pesquisa no Google, curtida ou comentário no Facebook, postagem no Twitter entre outras, pode determinar o perfil de um usuário e obter conhecimento sobre o seu gosto, entre outras diversas informações.

Com posse destas informações as empresas podem definir qual o seu público alvo e prever o que eles esperam de seus produtos. Com base em número de curtidas em determinada região, ou número de pesquisas relacionadas a seus produtos ou semelhantes, podem também ajudar empresas a escolherem o local mais adequado para abertura de uma nova filial, visando a escolha de um local que possa resultar em mais vendas e um lucro maior para a mesma. Essas informações providas através de redes sociais ou de outras ferramentas da internet podem manter estas empresas atualizadas sobre o contentamento ou não de seus clientes com seus produtos.

Analisando no que está acontecendo na internet, as empresas podem também ficar de olho em seus concorrentes, estando sempre atualizadas sobre o que eles estão publicando, suas promoções e a maneira com que interagem com seus clientes. Esse tipo de informação

pode se tornar uma oportunidade de mercado, uma vez que é possível identificar os pontos em que seus concorrentes deixam a desejar e os motivos pelo qual acabam perdendo uma venda.

Uma das melhores maneiras de melhorar os serviços ou produtos de uma empresa é conhecer a opinião dos clientes. Com este monitoramento de postagens realizado em tempo real, as empresas se mantêm atualizadas sobre o que os clientes estão falando sobre a sua marca e seus produtos.

Como dito anteriormente, para este trabalho foram coletadas informações no Twitter e Facebook, sendo a coleta no Twitter realizada por postagens que continham o termo “Fiat Mobi” e no Facebook coletando postagens e comentários na página “Fiat Automóveis Brasil”.

Foram coletados mais de sessenta e cinco mil registros divididos entre publicações e comentários postados nas duas redes sociais, no período de primeiro de janeiro de dois mil e dezesseis até a vinte e dois de maio do mesmo ano.

Realizando uma análise nas informações coletadas, foi possível identificar em alguns casos, postagens que mostram a satisfação dos clientes frente ao produto, foram encontradas publicações que se referiam ao Fiat Mobi de forma positiva e também de forma negativa (as positivas com um volume maior). Essa análise foi baseada em uma filtragem por palavras-chaves, onde foram definidas algumas palavras consideradas como sendo negativas e outras como positivas.

A Listagem 5 mostra código utilizado para filtrar as publicações consideradas como negativas, onde foi realizada uma consulta LINQ no objeto “postagensColetada” que continha todas as publicações coletadas em ambas as redes sociais. A linha 2 mostra o filtro principal da consulta, onde nele são filtradas todas as publicações que continham o termo “Mobi”, para assim retornar apenas as que faziam menção a este automóvel. Da linha 3 a linha 9 são passados os demais filtros, nesses a filtragem é realizada considerando apenas as palavras classificadas como chaves para publicações de conteúdo negativo. Assim esse método retorna apenas as publicações que se referem ao Fiat Mobi de forma não satisfatória.

Listagem 5 – Filtrando postagens com conteúdo negativo.

1. `var` postagensNeg = postagensColetadas.
2. `Where(x => x.Mensagem.Contains("Mobi"))`
3. `&& (x.Mensagem.Contains("caro ")`
4. `|| x.Mensagem.Contains("ruim ")`
5. `|| x.Mensagem.Contains("feio")`
6. `|| x.Mensagem.Contains("lixo ")`
7. `|| x.Mensagem.Contains("fraco ")`

```

8.             || x.Mensagem.Contains("apertado ")
9.             || x.Mensagem.Contains("pequeno ")
10.          )
11.      ).ToList();

```

Já na Listagem 6 é possível visualizar o código utilizado para filtrar publicações com conteúdo considerado positivo frente ao automóvel, onde também foi realizada uma consulta LINQ no objeto “postagensColetada”. Essa consulta é realizada de maneira semelhante a explicada anteriormente, onde na linha 2 mostra o filtro principal é realizada uma busca de todas as publicações que contém o termo “Mobi”, para assim realizar a filtragem das palavras-chaves apenas nestas publicações. Da linha 3 a linha 9 são passadas como filtro as palavras consideradas como chaves para publicações de conteúdo positivos, para assim apenas as publicações que se referem ao Fiat Mobi de forma positiva.

Listagem 6 – Filtrando postagens com conteúdo positivo.

```

1.  var postagensNeg = postagensColetadas.
2.      Where(x => x.Mensagem.Contains("Mobi")
3.          && (x.Mensagem.Contains("bom")
4.              || x.Mensagem.Contains("bonito")
5.              || x.Mensagem.Contains("lindo")
6.              || x.Mensagem.Contains("econômico")
7.              || x.Mensagem.Contains("gostei")
8.              || x.Mensagem.Contains("quero um")
9.              || x.Mensagem.Contains("amei")
10.         )
11.      ).ToList();

```

As postagens classificadas como negativas, em sua maioria referiam-se à insatisfação dos clientes com o preço e tamanho do espaço interno do carro, a Figura 15 mostra algumas destas publicações. Já as publicações que se referem de forma positiva, mostram a satisfação com o automóvel, sua economia e beleza, assim como pode ser visto em algumas das publicações coletadas visíveis na Figura 16.

Figura 15 - Publicações negativas referente ao Fiat Mobi.

Hoje eu fui ver este carro na concessionária e: versão rabeala por 33 mil e a versão completinha 46 mil kkkkkkk meu eu ri..... o carro realmente é pobi ao inv
Lixooooo prefiro meu UP!!! Mobi cópia mal feita!!!!
Mobi e up as carroças mãos feio do Brasil e carroça nacional, isso é Brasil
Oh povo feio esse MOBI
Mobi lixo
Ainda não entendi a que veio o Mobi. É um Uno mais simples, mais apertado, sem inovação tecnológica alguma sobre o irmão mais velho... E o preço aproxir
Verdade seja dita, a FIAT pecou feio em não lançar o MOBI com motor de 3 cilindros, quem adorou foi a PEUGEOT que em Julho lançará o 208 com 3 cilind
Só brasileiro burro e idiota para pagar 32.000,00 por um playmobil básico pelado e ultrapassado, essas montadoras de carros no Brasil, são iguais aos políticos
O Mobi vem com motor Evo-Fire, 1.0 e 1.4 desenvolvidos para o novo uno, a partir do motor Fire. São motores ultrapassados, em vista dos motores de 3 cil
Não vejo nenhuma novidade no Fiat MOBI. É uma Fiat UNO com frente do FREEMONT, com o mesmo Motor manco e caro.
Mobi-cho feio
MOBICA carro caro demais sai fora!!
MOBIcho feio!
Fiat Mobilete econômico como a cinquentinha, caro como carro de luxo. Adquiria já o seu Fiat mobilete nas concessionárias!
Olhem os comentários contra o Mobi! Vejam os perfis! Um se diz funcionário da GM, outro, é um pia que nem título tem (16 anos)! Vão rachar uma lenha! A F
A fiartinha que ter vergonha de apresentar um carro tão ruim como Mobi é caro d+ antes reutilizassem o uno com com desenho do Mobi
Adoro a Fiat, mas esse mobi TÁ MUITO CARO POR NADA!! O CARRO NÃO É ISSO TUDO PRA CUSTAR TANTO
Acabei de voltar de viagem, estava em Portugal, lá tem um carro da opel chamado karl, carro completo e custa 11.990 euros. Aqui no Brasil só os otarios gas
Mobi, mobilidade, se for pra andar apertado pego um Bus
Não achei o verdadeiro sentido da proosta mobi, pagar 31k na versão de entrada e carrear 2 pessoas atrás apertado enquanto com esse valor da pra con

Figura 16 - Publicações positivas referente ao Fiat Mobi.

tive quando saiu 2010/2011 lindo demais só vendi porque clonaram minha placa, se não tava com ele ate hoje, curto hoje uma Mercedes bom demais! A FIAT podia cri
Lindooooo!!!!!! Eu quero um Mobi azul, este caminho é uma delícia de dirigir 🚗🚗🚗🚗
Gostei muito desse Mobi da Fiat
Lindo quero um mobi vermelho,
quero um mobi
Uau como é lindo estes robôs trabalhando na montagem do mobi lindo di mais. Incrível aonde a magia acontece ao vivo parabéns Fiat o lugar ai é lindo
Eu tinha um ideia maravilhoso divino e agora troquei por um Mobi zerinho, ótimo carro, claro q troquei devido a economia pq por mim nunca mais saia do Idea, nunca tive pr
Q lindooooooooooooooooo. Ainda vou buscar o meu Mobi azul e vários outros Fiat's pela vida. Eu quero um Mobi azul, fica é topp de mais, adorei fazer um teste drive n
gostei mais do video, mas prefiro ter uma toro do que um mobi
Amei esse mobi !!! muito lindo !!!!!
Eu quero um novo MOBI da FIAT the best
Esperacei o mobi sedan 1,4 e lindo d mais
Lindo Mobi completo vermelho♥
Comprei um Mobi e não me arrependo. É lindo, super econômico. O design encanta a todos. Recomendando.
Lindo lindo o mobi show.
Eu quero um Mobi azul !!!!!. É muitooooooooo lindoooo, eu andei num Mobi fazendo teste drive e gostei ainda mais. DH uma delícia, confortável, acabamento interno n
Tenho um Up e recentemente minha esposa adquiriu um Mobi. Sem dúvida o UP é superior em toda a questão mecânica, mas o apelo visual mais moderno vai fazer muit
Quero um desse ... o meu e o novo uno mas já quero um mobi
Quero ganhar meu Fiat Mobi no dia dos namorados! Kkk

Foram encontradas também, publicações que comparavam o Fiat Mobi com automóveis de outra marca, questionando outros clientes e até mesmo a marca, de qual seria o motivo para adquirir este ao invés de algum carro de outra marca.

Todas essas publicações funcionam como um boca a boca em grande proporção, podendo influenciar diretamente nas vendas. Empresas que provem destas informações em tempo real, podem agir de tal forma que minimize a influência negativa que este “boca a boca

on-line” pode gerar, ou até mesmo encontrar alguma forma de melhorar e evoluir seu produto para ampliar as vendas.

Analisando as publicações coletadas foi possível identificar também que a maioria das interações dos usuários com a Fiat, nas redes sociais, ocorrem entre a quinta-feira e o sábado, nos horários entre as dezenove horas e as vinte e duas horas, assim como mostram as Figuras 17 e 18.

Para chegar a estas informações, foram realizadas consultas no objeto que continha todas as publicações coletada. A Listagem 7 mostra a consulta que retorna o número de publicações por dia da semana, para esse método também foi realizada uma consulta LINQ onde as publicações foram agrupadas por data, considerando apenas o dia da semana em que foram publicadas, retornando então o número de publicações para cada dia.

Já nas Listagem 8 é possível enxergar o método utilizado para consultar o número de publicações por hora do dia, para esse método também foi realizada uma consulta LINQ, onde os dados foram agrupados por hora de publicação, assim retornando o número de publicações por hora do dia.

Listagem 7 – Consultando o número de postagens por dia da semana.

```

1.  var postagensPorDia = (from post in postagensColetadas
2.      group post by post.DataCriacao.DayOfWeek
3.      into dados
4.      select new
5.          {
6.              DiaSemana = dados.Key,
7.              Count = dados.Count(),
8.          }).ToList();

```

Listagem 8 – Consultando o número de postagens por hora do dia.

```

1.  var postagensPorDia = (from post in postagensColetadas
2.      group post by post.DataCriacao.Hour
3.      into dados
4.      select new
5.          {
6.              DiaSemana = dados.Key,
7.              Count = dados.Count(),
8.          }).ToList();

```

Figura 17 - Gráfico de publicações por dia da semana.

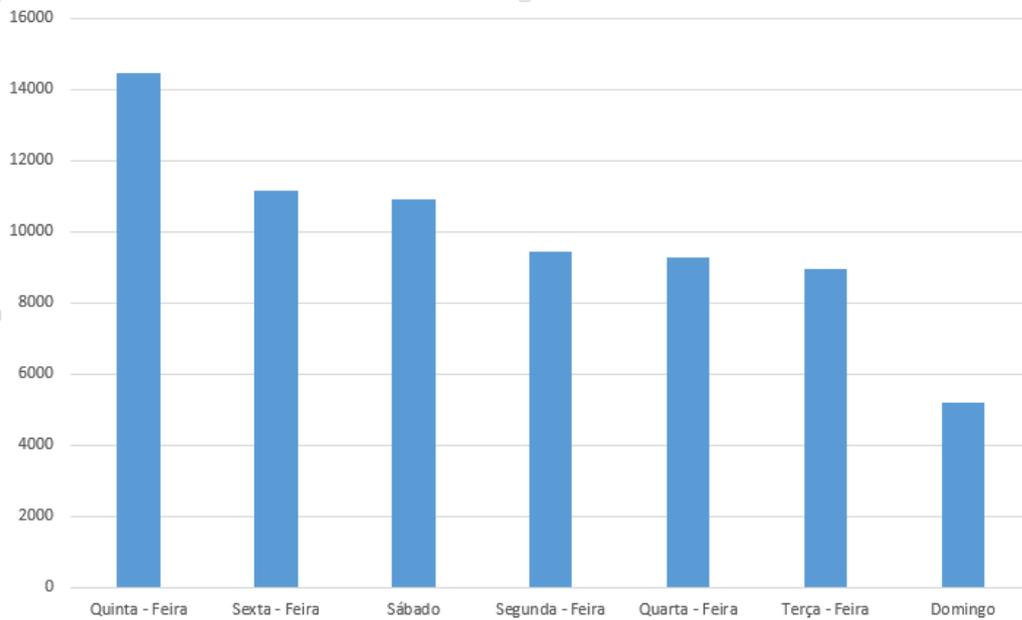
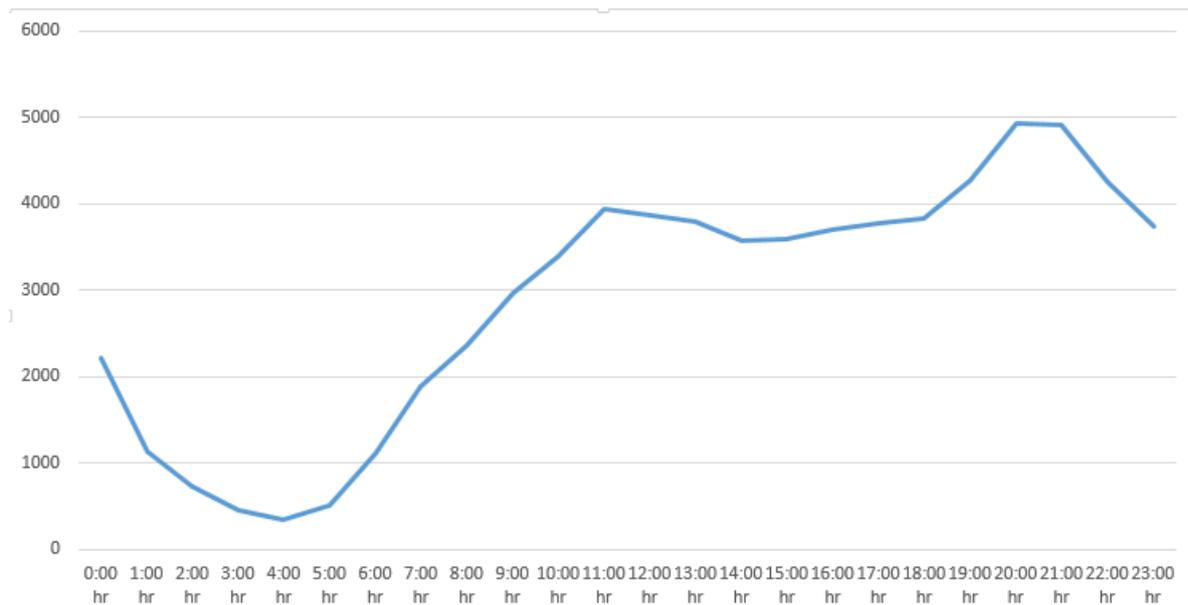


Figura 18 - Gráfico de publicações por hora.



Esse tipo de informação pode ser importante para empresas, ajudando a definir qual o melhor horário para divulgar promoções e produtos, ou até mesmo em que momento deve-se

dar uma atenção maior as redes sociais para ficar de olho sobre o que está sendo publicado, sendo sobre o seu produto ou até mesmo do concorrente.

5 CONCLUSÃO

Para Siqueira (2010) o monitoramento de redes sociais é importante ver qual é a situação da sua empresa do ponto de vista do cliente. Ouvir se eles estão falando sobre a marca e o que é que estão falando, se sua empresa e seus produtos estão agradando ou não. Assim, pode-se identificar quais estratégias estão funcionando bem e quais ainda precisam ser melhoradas.

Ter controle e poder de análise em tempo real sobre dados é o que pode diferenciar uma empresa. Com o apoio de ferramentas tecnológicas para análise e classificação é possível interpretar e gerar informações importantes sobre um grande volume de dados, informações estas que contribuam para a tomada de decisão.

Com as ferramentas adequadas o Big Data oferece informações relevantes que auxiliam na tomada de decisões, baseado no conceito 5V (Valor, Veracidade, Variedade, Volume e Velocidade) Big Data oferece informações muito próximas da realidade por permitir análise em um grande volume de dados.

Big Data influencia o processo de tomada de decisões, pois garante um grande acervo de informações a serem tratadas, onde estudos podem ser efetuados para melhor avaliar um futuro serviço a se oferecer ou uma provável demanda, além de permitir avaliar também os produtos que já se encontram no mercado analisando dados sobre estes disponíveis na internet, estabelecendo um feedback em tempo real. Utilizando-se dessa ferramenta pode se estabelecer estratégias que visam garantir um possível negócio ou qualquer outro serviço e identificar os possíveis compradores (SILVA, 2014).

Com este trabalho foi possível verificar os benefícios de análise de dados aplicada em redes sociais, mostrando que essa análise em um grande volume de dados pode prover informações, em tempo real, que auxiliem empresas na tomada de decisões.

REFERÊNCIAS

- ADAMI, Anna. **Big Data**. 2011. Disponível em: <http://goo.gl/SZ3eq>. Acesso em: 12 nov. 2015.
- ALECRIM, Emerson. **O que é Big Data?** 2013. Disponível em: <http://goo.gl/h6PPUR>. Acesso em: 12 nov. 2015.
- BORBA, D. K.; SANTOS, L. F. A; KAWAMOTO, L. T. **Big Data: percepção dos usuários sobre vantagens e invasão de privacidade**. Universidade Centro Paula Souza. São Paulo, SP, Brasil. 2013. Disponível em: <http://goo.gl/vLN6Rr>. Acessado em: 11 out. 2015.
- BRITO, Ricardo W. **Bancos de Dados NoSQL x SGBDs Relacionais: Análise Comparativa**. Universidade de Fortaleza. Fortaleza, CE, Brasil. 2011. Disponível em: <http://goo.gl/nzcwzO>. Acessado em: 05 mar. 2016.
- CISCO. Homepage da Instituição. Disponível em: <http://goo.gl/vl3nif>. Acesso em: 11 out. 2015.
- COMSCHOOL. **As 5 principais redes sociais mais usadas no Brasil**. 2015. Disponível em <http://goo.gl/OhfMhc>. Acessado em: 24 mar. 2016.
- DAMASCENO, Sergio. **As tendências de web, mobile e social em 2015**. 2015 . Disponível em: <http://goo.gl/y2sHX5>. Acessado em: 24 abr. 2016.
- DAVENPORT, Thomas H.; KIM, Junio. **Dados Demais!** Elsevier Editora, Rio de Janeiro, 2014.
- DELSOTO, Diego. **A influência do Big Data no Business Intelligence**. Disponível em: <http://goo.gl/10PtPt>. Acesso em: 12 nov. 2015.
- DIANA, Mauricio; GEROSA, Marco Aurélio. **NOSQL na Web 2.0: Um Estudo Comparativo de Bancos Não-Relacionais para Armazenamento de Dados na Web 2.0**. Universidade de São Paulo. São Paulo. 2010. Disponível em: <http://goo.gl/CtAaoO>. Acesso em: 20 nov. 2015.
- FOWLER, Martin. **NoSQL Essencial: Um Guia Conciso para o Mundo Emergente da Persistência Poliglota**. Novatec Editora, São Paulo, 2013.

- FRANÇA, T. C.; FARIA, F. F.; RANGEL, F. M.; FARIAS, C. M.; OLIVEIRA, J. **Big Social Data: Princípios sobre Coleta, Tratamento e Análise de Dados Sociais**. Universidade Federal do Paraná, 2014. Disponível em: <http://goo.gl/HveAwa>. Acessado em: 31 out. 2015.
- IBM Corporation; **Conceito:** Orientação a Objetos e Bancos de Dados Relacionais. Disponível em <http://goo.gl/hvY9w9>. Acessado em: 27 mar. 2016.
- LEITE, Gleidson Sobreira. **Análise Comparativa do Teorema CAP Entre Bancos de Dados NoSQL e Bancos de Dados Relacionais**. Faculdade Farias Brito, Fortaleza, 2010. Disponível em: <http://goo.gl/DXqS9r>. Acessado em: 27 mar. 2016.
- LIMA, Fábriço. **Lista com um ranking de popularidade de SGBDs**. 2013. Disponível em: <http://goo.gl/0j8uow>. Acesso em: 26 mar. 2016.
- LÓSCIO, Bernadette F.; OLIVEIRA, Hélio Rodrigues; PONTES, Jonas C. de Sousa. **NoSQL no desenvolvimento de aplicações Web colaborativas**. VIII SIMPÓSIO BRASILEIRO DE SISTEMAS COLABORATIVOS. 2011. Disponível em: <http://goo.gl/zXl1jrm>. Acesso em: 20 nov. 2015.
- MACORATTI, José Carlos. **ASP. NET - Apresentando LINQ - Language Integrated Query**. 2012. Disponível em: <http://goo.gl/thC8R2>. Acessado em: 04 jun. 2016.
- MAYO, Joe. **Documentation – Linq to Twitter**. 2015. Disponível em: <https://goo.gl/sWW7tF>. Acessado em: 04 jun. 2016.
- NETO, Franciso G. P. **Banco de Dados Relacional – Exemplo**, 2008. Disponível em: <http://goo.gl/8YF51v>. Acesso em: 20 nov. 2015.
- SANTANA, Otávio Gonçalves. **Introdução ao Conceito de Big Data**. 2013. Disponível em: <http://goo.gl/EL0FRO>. Acesso em: 20 nov. 2015.
- SCHÖNBERGER, Viktor.; CUKIER, Kenneth. **Big Data: Como extrair volume, variedade, velocidade e valor da avalanche de informação cotidiana**. Elsevier Editora, Rio de Janeiro, 2013.
- SILVA, Andreiver Mateus Ferreira. **Big Data e a Tomada de Decisão**. 2014. Disponível em: <http://goo.gl/DLmfEv>. Acessado em: 04 jun. 2016.
- SIQUEIRA, André. **8 razões para monitorar a sua marca na Internet**, 2010. Disponível em: <http://goo.gl/9WuoaL>. Acessado em: 28 mai. 2016.
- TAURION, Cezar. **Big Data**. Editora Brasport, Rio de Janeiro, 2013. Disponível em: <https://goo.gl/34H4sp>. Acessado em: 16 out. 2015.