Soluções de BI 2.0 para Análise de Dados a partir do Twitter®: Eleições 2014

Jonatas A. Tavares[1], Flávio Ceci[1,2]

[1] Pós-Graduação em Engenharia de Projeto de Software. Universidade do Sul de Santa Catarina (UNISUL) – Palhoça – SC – Brasil

[2] Departamento de Engenharia do Conhecimento – Universidade Federal de Santa Catarina (UFSC) – Florianópolis – SC – Brasil
jonatasatavares@gmail.com, flavio.ceci@unisul.br

Abstract. According to its growing trend, the Internet has become a rich information environment, containing many kinds of matters. By consequence, this environment has become very attractive to organizations. However, these organizations need to acquire tools that enable proper treatment of this information. Given this context, this article has as solution proposal, present a Business Intelligence 2.0 solution that allows the extraction of information in the social network Twitter, unstructured manner, perform the ETL process, and insert it into a database, enabling a graphical analysis of information.

Resumo. De acordo com sua crescente evolução, a internet transformou-se em um ambiente rico de informações, contendo diversos tipos de assuntos. Por consequência disto, este ambiente tornou-se muito atrativo às organizações. Contudo, estas organizações necessitam adquirir ferramentas que possibilitem o devido tratamento destas informações. Diante deste contexto, este artigo procura apresentar uma solução de Business Intelligence 2.0 que permita a extração de informações contidas na rede social Twitter®, de maneira não estruturada, realize o processo de ETL e as insira em uma base de dados, possibilitando uma análise gráfica das informações obtidas.

1.Introdução

A utilização de soluções de *Business Intelligence* (BI) está cada vez mais presente nas organizações. O uso dessas soluções tornou-se mais frequente, devido a sua abrangência, pois podem ser utilizadas em qualquer área de atuação das organizações, tendo como foco o auxílio nas tomadas de decisões.

O mercado, por sua vez, torna-se mais competitivo. São inúmeras as empresas de diversos tipos de segmentos, contribuindo para uma melhoria nos serviços oferecidos e na busca pela satisfação dos clientes. Nesta busca, nota-se que a satisfação do cliente só é alcançada quando se atinge a sua necessidade de negócio.

Com isso, as organizações iniciaram uma procura por ferramentas com o intuito de conhecer melhor o ambiente de negócio dos seus clientes e adquirir vantagem competitiva em relação às demais organizações, servindo assim de amparo no processo de tomada de decisão.

O processo de tomada de decisão é de extrema importância para as organizações, pois o futuro destas depende das ações tomadas. Contudo, o processo de decisão não pode ser realizado de forma simplista, existe uma necessidade de analisar criteriosamente as informações disponíveis.

Além desta análise criteriosa, outra dificuldade também foi encontrada pelas organizações durante a execução dos processos decisórios. Devido ao crescimento demasiado das informações no ambiente externo, principalmente na Web, se torna difícil identificar quais informações são realmente necessárias resultando num alto grau de complexidade da análise. Tendo em vista o alto grau de complexidade, as organizações identificaram que haveria a necessidade de utilizar um ferramental tecnológico capaz de processar esta grande quantidade de informações.

Surgiram, assim, ferramentas que são capazes de extrair esta grande quantidade de informações da WEB, armazenar os dados relevantes em estruturas de banco de dados organizadas e processá-los, gerando conhecimento para auxiliar nos processos decisórios das organizações. Características que resultaram em um novo conceito, o BI 2.0.

Na visão de Pintas e Siqueira (2011), a maior deficiência das soluções tradicionais de BI está na latência entre o acontecimento do evento e a tomada de decisão. Segundo os autores em questão, o BI 2.0 tem como foco atacar essa latência.

Este artigo tem como finalidade demonstrar uma solução de BI 2.0 que permita usar as informações contidas no *Twitter®* e inseri-las em um repositório de dados, a fim de apoiar uma decisão futura.

Nas próximas seções estão descritos os conceitos e metodologias de assuntos relacionados ao contexto do artigo, proposta de solução contendo uma figura demonstrativa do experimento em si, o próprio experimento com informações detalhadas da solução e por fim as considerações finais realizando um desfecho do artigo e sugestões para pesquisas futuras.

2.Referencial Teórico

Esta seção tem como principal objetivo apresentar um referencial bibliográfico para amparar os temas e assuntos abordados no artigo, além de auxiliar no seu desenvolvimento.

2.1. Business Intelligence (BI)

De acordo com Turban *et al*. (2009), o ambiente de negócios no qual as empresas operam atualmente está se tornando cada vez mais complexo e mutante. O volume de informações armazenadas nos bancos de dados das organizações é cada vez maior. Contudo, apenas o armazenamento dessas informações não é suficiente, há a necessidade de analisá-las e utilizá-las de forma inteligente na tomada de decisões das empresas.

Segundo Côrtes (2002), *Business Intelligence* é um conjunto de conceitos e metodologias que visa ao apoio à tomada de decisões nos negócios a partir da transformação do dado em informação e da informação em conhecimento.

2.2.Business Intelligence 2.0

Segundo Martins (2008), o propósito dos ambientes de BI 2.0 é melhorar o desempenho dos processos de tomada de decisão, reduzindo o tempo entre a ocorrência de um evento no ambiente transacional e o momento quando uma decisão é tomada no ambiente informacional.

Nelson (2010), por sua vez, ensina que BI 2.0 implica um "afastamento do armazém de dados padrão que as ferramentas de inteligência de negócios têm usado" e que "dará lugar ao contexto e a necessidade de relacionar informações de forma rápida a partir de muitas fontes".

2.3.Extração de Informação

Um dos problemas que as organizações têm enfrentado para trabalhar com o conhecimento é como encontrá-lo, recuperá-lo, armazená-lo, e compartilhá-lo entre os seus membros (CECI, 2010). Como o foco deste artigo são as redes sociais, as quais contemplam informações que podem auxiliar na tomada de decisão, se faz necessária a utilização de recursos para extração desses dados necessários para a organização.

2.4.Reconhecimento de Entidades Nomeadas (NER)

Para Ceci, Pietrobon e Gonçalves (2012), Reconhecimento de Entidades Nomeadas (NER) é considerado uma parte da extração de informações, onde o objetivo é encontrar e categorizar seções de texto em categorias pré-estabelecidas.

Ceci (2012) explica que o NER (*Named Entity Recognition*) é uma técnica que tem como objetivo encontrar as "fronteiras" de um termo no texto e, se disponível uma base de conhecimento, também classificar este termo, como, por exemplo, pode-se reconhecer o termo "UFSC" e apresentá-lo como uma organização.

2.5.Descoberta de Conhecimento em Texto (KDT)

O processo KDT é definido como a extração de padrões relevantes e não triviais a partir de bases de dados semi ou não estruturadas. Também, utiliza técnicas da mineração de dados, mas, nesse caso, usam-se técnicas de processamento de linguagem natural para extrair conceitos de texto e mais uma vez análises estatísticas, mas para recuperar padrões e técnicas de visualização, permitindo análises interativas (GONÇALVES, 2006).

Mooney e Nahm (2003) completam que é um processo para encontrar padrões interessantes e úteis, modelos, direções, tendências ou regras a partir de textos não estruturados.

2.6.**WEB 2.0**

De acordo com Bressan (2007), em linhas gerais, Web 2.0 diria respeito a uma segunda geração de serviços e aplicativos da rede e a recursos, tecnologias e conceitos que permitem um maior grau de interatividade e colaboração na utilização da Internet.

Segundo Moura (2012), o termo Web 2.0 está associado a aplicações Web, em que o objetivo principal é facilitar os seguintes aspectos: compartilhamento de

informações de maneira interativa, interoperabilidade, desenvolvimento com foco no usuário e colaboração na World Wide Web (WWW).

3.Metodologia

Nesta seção estão descritos e ilustrados o método e a proposta de solução contendo as ferramentas utilizadas para o seu desenvolvimento. Para facilitar o entendimento foi criado uma figura, figura 1, que ilustra todo o detalhamento de cada componente existente na proposta de solução.

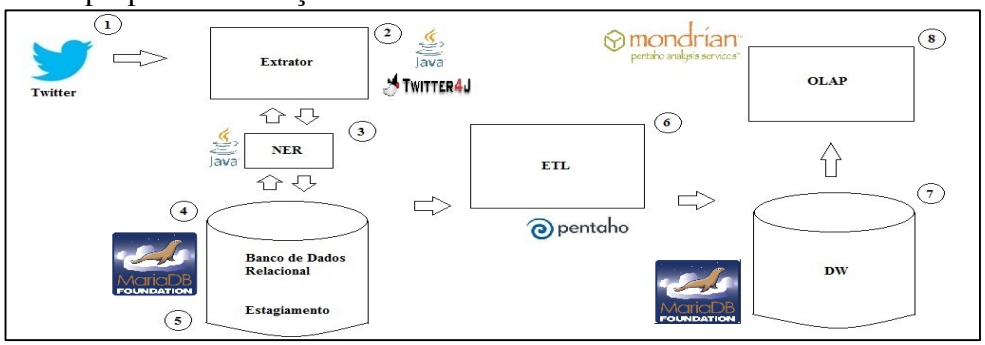


Figura 1 – Proposta de Solução

1-**Twitter**: Rede social utilizada na proposta de solução que contempla os dados não estruturados essenciais para a população da base de dados.
2-**Extrator**: Ferramenta que faz a extração dos *tweets* de acordo com os termos de interesse passados pelo usuário. Para desenvolvimento do extrator foi utilizada a linguagem de programação Java com a API Twitter4j.
3-**NER**: Algoritmo responsável por identificar os termos e entidades nos dados coletados (*tweets*). Para desenvolvimento do NER foi utilizado a linguagem de programação Java.
4-**Banco de Dados Relacional**: Estrutura de banco de dados que contempla os dados coletados através da ferramenta Extrator. Para o desenvolvimento do banco de dados foi utilizado a ferramenta MariaDB.
5-**Estagiamento**: Estrutura de banco de dados que contempla o estruturado os relacionamentos entre os termos, entidades e os dados coletados. No modelo de estagiamento também foi utilizado o MariaDB.
6-**ETL**: Ferramenta utilizada para estruturar os dados no banco de dados dimensional. Para o desenvolvimento da ETL foi utilizado o Pentaho.
7-**DW (*Data Warehouse*)**: Estrutura de banco de dados que contempla as dimensões e a tabela fato essenciais para permitir e agilizar as consultas. Para o desenvolvimento do DW foi utilizado o MariaDB.
8-**OLAP**: Ferramenta utilizada para realizar as consultas no DW. Para o desenvolvimento foi utilizado a ferramenta Mondrian.

4.Análise

Nesta seção está demonstrado o detalhamento da solução de BI 2.0 criada, os modelos de dados e os resultados dos testes realizados no experimento. Esta seção também

contempla as tecnologias utilizadas, explicação do algoritmo de coleta de tweets, informações sobre estrutura do banco de dados e demais informações sobre OLAP.

4.1.Fonte de Dados

Para o desenvolvimento da proposta de solução, primeiramente, foi necessário identificar a fonte de dados utilizada para coleta de informações que satisfizessem a análise. Como este trabalho trata-se de uma solução de BI 2.0 foi necessário identificar uma fonte de dados contida na Web 2.0. A fonte de dados escolhida foi o Twitter®.

O Twitter® é caracterizado por ser um microblog, ou seja, um pequeno blog. Segundo O' Reilly e Milstein (2009, p.13), o serviço de comunicação foi criado em março de 2006 pela Obvious e "inicia-se como um projeto sem grandes pretensões, idealizado por uma empresa de podcasting de São Francisco, e não demorou para se tornar o principal projeto dela".

4.2.Extrator de Dados

Para desenvolvimento do extrator de dados foi utilizado uma API do Twitter, Twitter4j, pois apresentou métodos e interfaces simples para desenvolvimento e coleta das informações. Para as coletas, foi necessária, ainda, a criação de uma conta e também um perfil.

Desta forma, o algoritmo do extrator utilizou métodos da própria API que tem como característica armazenar em uma lista palavras-chaves ou termos mais frequentes e para cada tweet disponível na rede social Twitter é realizado uma comparação de *string* identificando tweets que possuem uma dessas palavras existentes na lista.

O desenvolvimento do coletor contempla uma tela de *login*, necessária para autenticação dos usuários. A validação dos dados de *login* é realizada com a comparação entre os dados inseridos pelo usuário na tela de *login* e os dados já existentes no banco de dados. Caso a validação retorne um valor positivo a autenticação é realizada com sucesso e o usuário é redirecionado para a próxima tela, tela inicial, do sistema, caso negativo, o usuário é informado que a autenticação falhou.

A tela inicial contempla informações do *stream* de *tweets*, serviço responsável pelas coletas. Nas informações apresentadas nesta tela estão os status do serviço, podendo ser este "ativo" ou "inativo", e desde quando o serviço está neste status, ou seja, a última data e hora que houve a alteração de status. Nesta tela também é possível realizar a inicialização ou a pausa do serviço de *stream*.

Na tela seguinte, tela de cadastro de termos, é possível realizar o cadastro e remoção dos termos para pesquisa de *tweets*. Os termos são palavras, *hashtags* e perfis utilizados como buscas por *tweets* na API do Twitter®.

No módulo Coletor, ainda, existe uma tela que demonstra o último *tweet* coletado, com informações do usuário que realizou o cadastro do *tweet* e data. Esta tela foi criada apenas para confirmar o funcionamento do *stream* pela interface.

Para finalizar, a última tela do módulo Coletor, tela de carregamento de entidades, possui como finalidade o cadastramento de entidades. As entidades são todos os elementos utilizados no experimento para demonstrar e classificar os tipos de dados provenientes do *Data Warehouse*.

4.3.Banco de Dados

O modelo dimensional da proposta de solução é composto por dez tabelas, dentre as quais oito são tabelas de dimensão, uma é tabela fato e outra é uma tabela de estagiamento, sendo elas: dim_pessoa, dim_evento, dim_partido, dim_tempo, dim_termo, dim_assunto, dim_cargo, dim_semente, est_fato e fato_tcc. Segue figura para ilustrar o modelo dimensional proposto no sistema.

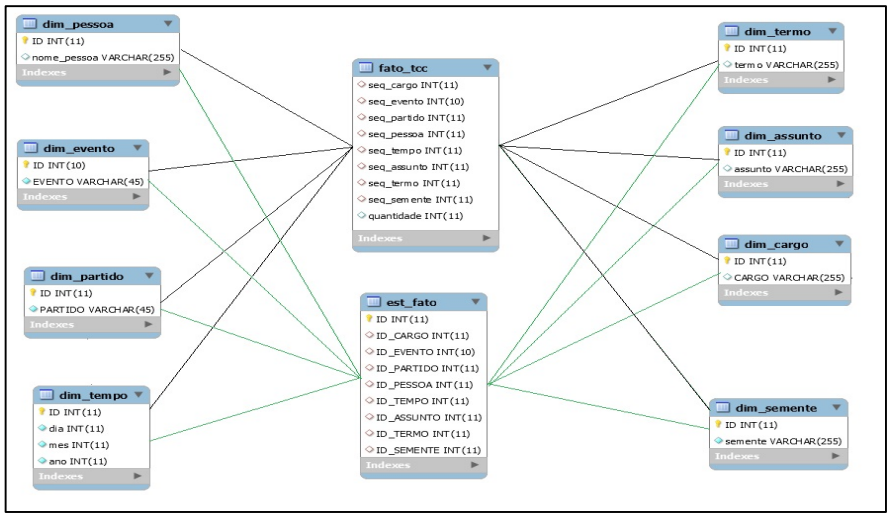


Figura 2 – Modelo Dimensional

A sessão a seguir demonstra o algoritmo de identificação de entidades.

4.4.Algoritmo de Identificação de Entidades

Para desenvolvimento do algoritmo de identificação de entidades foi utilizado a linguagem Java e técnicas de extração de termos e entidades (NER).

Nesta etapa, primeiramente, foi capturada uma amostra de 100 tweets coletados da rede social Twitter e identificado os termos mais frequentes e relevantes.

Com a lista de termos mais relevantes criada, esta lista por sua vez foi armazenada em uma tabela do banco de dados relacional.

Assim, o algoritmo percorre a tabela que contém o armazenamento dos tweets e para cada tweet é percorrido a tabela de listagem de termos frequentes. Em seguida, é feita uma comparação de string. Caso seja localizado algum termo no tweet, estes são armazenados em uma nova estrutura de banco de dados denominada estagiamento.

4.5.ETL

Conforme informação anterior, para criação do modelo dimensional foi utilizado o processo de ETL, extraindo informações do banco de estagiamento e carregando para estruturas organizadas com a finalidade de melhorar e facilitar as pesquisas.

Com o auxílio da ferramenta Pentaho Data Integration foram criadas *steps* e transformações com a finalidade de popular as dimensões do modelo dimensional e da tabela fato.

Esta carga possui como entrada de dados os *tweets* capturados e armazenados no modelo relacional. Com os dados dos *tweets*, foram criados mais dois *steps* para identificação das entidades e termos destes *tweets* e em seguida realizado um *merge*. Com isso, é feito a identificação das dimensões e os seus referidos dados e associados utilizando um identificador único entre a tabela fato e suas dimensões.

4.6.Ferramentas OLAP

Por último foi utilizada uma ferramenta OLAP, necessária para realizar pesquisas no banco de dados e trazer resultados em gráficos e relatórios para serem analisados no processo decisório.

A ferramenta utilizada foi o Mondrian. Nesta ferramenta foram criadas estruturas de dados, denominadas cubos. Os cubos têm como finalidade criar estruturas organizadas para facilitar as consultas no modelo de banco de dados dimensional.

Por fim, as consultas criadas são processadas e os resultados demonstrados em uma ferramenta de relatórios. Para esta tarefa, a ferramenta utilizada foi o Pentaho.

5. Discussão dos Resultados

Esta seção apresenta a avaliação efetivada contemplando a importância da ferramenta para apoio nas decisões de temas cotidianos. Nesta proposta o tema utilizado para análise foi as eleições 2014, contudo a ferramenta pode ser utilizada para os demais temas cotidianos desde que sejam realizados alguns ajustes no módulo de análise.

Estão apresentados alguns exemplos de gráficos gerados pela ferramenta e uma breve interpretação destes.

A primeira análise demonstrada é a quantidade total de *tweets* coletados relacionados aos candidatos à presidência. Este gráfico é importante para identificar quais candidatos foram mais comentados e foram mais expressivos durante o período da análise. Segue figura que ilustra este gráfico.

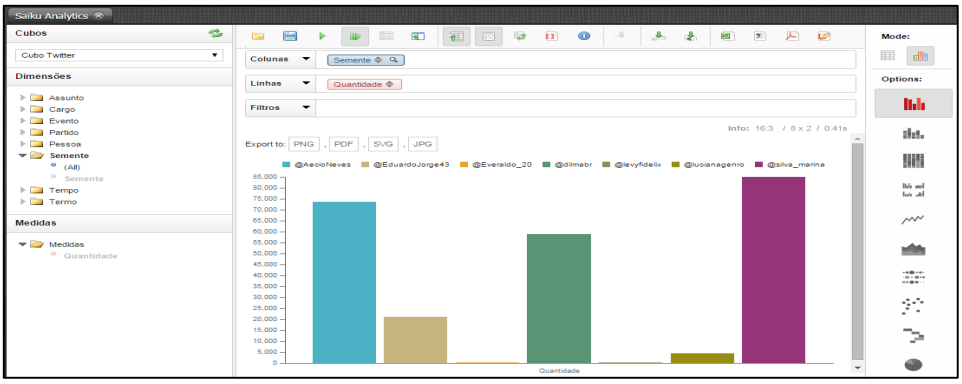


Figura 4 – Total Tweets Candidatos Presidência

Através deste gráfico verifica-se que os candidatos à presidência das eleições de 2014 com mais citações nos *tweets* foram os candidatos Marina Silva, Aécio Neves e Dilma Roussef. Já os candidatos com menor número de citações nos *tweets* foram Pastor Everaldo e Levy Fidelix com cerca de 350 *tweets*. A interpretação deste gráfico

também foi interessante pela definição do grau de importância dos candidatos em relação aos eleitores. A concentração de comentários e críticas, sendo positivas ou não a um candidato demonstrou que a população se interessou pelos assuntos abordados pelo candidato. Prova disso, foi que um dos três candidatos com maiores citações foi eleito para a presidência.

No próximo gráfico está demonstrada a quantidade de citações nos *tweets*, separados por assuntos, de cada candidato à presidência das eleições de 2014.

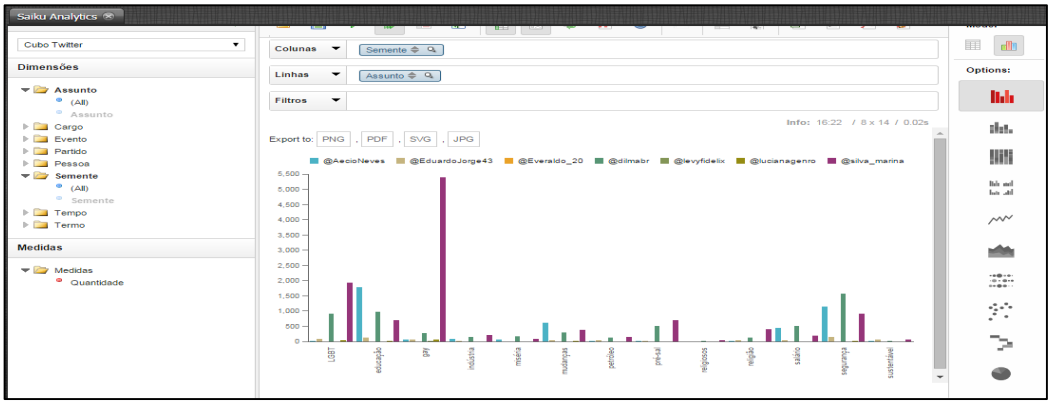
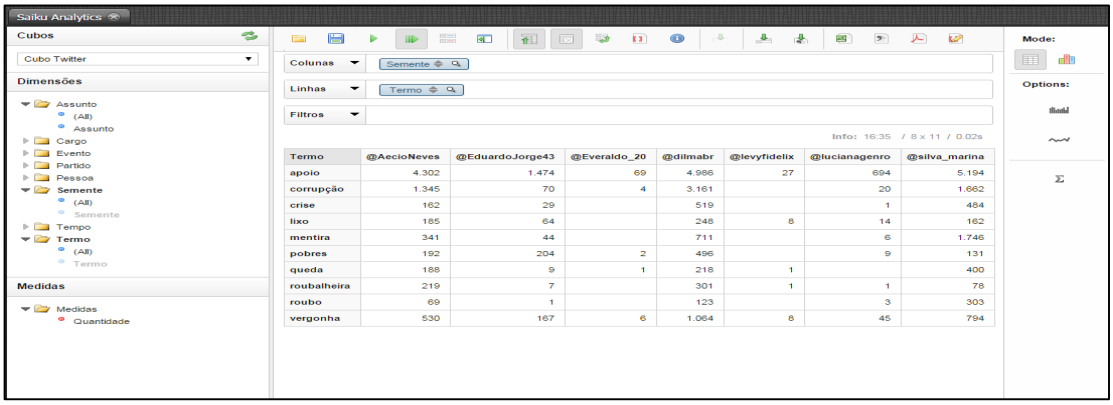


Figura 6 – Total Tweets Presidência Assunto

Analisando o gráfico, verifica-se que a candidata Marina Silva foi a mais citada nos *tweets* em alguns assuntos, sendo eles: LGBT, gay, pré-sal, indústria e religião. A candidata Dilma por sua vez foi a mais citada em relação a salário e segurança. Já Aécio Neves foi o mais citado em educação e mudanças.

Se observarmos e traçarmos um comparativo com os assuntos tratados no cenário das eleições e no cenário atual, verificamos que são assuntos ainda discutidos, demonstrando também o grau de importância. Por exemplo, referente à questão de segurança pública, recentemente foi discutido uma PEC (Proposta de Ementa de Constituição) para a redução da maioridade penal. Em relação ao assunto salário, recentemente, também foi criado um Programa de Proteção ao Emprego com definições sobre jornada de trabalho e salário dos empregados.

A próxima imagem representa a quantidade de citações de termos frequentes apresentados nos *tweets* em relação aos candidatos à presidência das eleições de 2014.



Termo	@AecioNeves	@EduardoJorge43	@Everaldo_20	@dilmabr	@levyfidelix	@lucianagenro	@silva_marina
apoio	4.302	1.474	69	4.986	27	694	5.194
corrupção	1.345	70	4	3.161		20	1.662
crise	162	29		519		1	484
lixo	185	64		248	8	14	162
mentira	341	44		711		6	1.746
pobres	192	204	2	496		9	131
queda	188	9	1	218	1		400
roubalheira	219	7		301	1	1	78
roubo	69	1		123		3	303
vergonha	530	167	6	1.064	8	45	794

Figura 7 – Total Tweets Termo

Neste gráfico, verificou-se que a maior quantidade de termos frequentes relacionados à candidata Marina Silva, foram relacionados aos termos apoio, mentira, queda e roubo. Já em relação à candidata Dilma, foram os termos corrupção, pobres, roubalheira e vergonha. O candidato Aécio Neves também teve um número expressivo em relação ao termo apoio.

Em nova análise comparativa com o cenário atual, podemos identificar também que estes termos estão sendo tratados com grande frequência. Por exemplo, a questão da corrupção com os escândalos da operação Lava Jato. Este termo também apareceu recentemente junto com assuntos relacionados à CBF (Confederação Brasileira de Futebol).

Estes gráficos demonstrados nesta seção são alguns exemplos de análise realizadas com o intuito de demonstrar a importância deste tipo de ferramenta, ou ainda, deste tipo de análise. Deve ser lembrado que a ferramenta é necessária para apoio à tomada de decisão, com isso outras análises devem ser feitas para chegar a um resultado final.

6.Conclusões

Em análise aos ambientes corporativos atuais, observa-se cada vez mais um acúmulo demasiado de informações, podendo elas ter origem interna ou até mesmo externa das organizações, como no caso foi citado neste trabalho a Web. Contudo o acúmulo demasiado de informações pode causar problemas para a própria organização no processo de tomada de decisão. O grande problema está na forma em que esta grande quantidade de informação é coletada e processada, pois informações que não são úteis para o processo decisório, poderão afetar no seu resultado final. Com isso, as organizações têm a necessidade de utilizar ferramentas com a finalidade de auxiliar tanto nas tarefas diárias quanto nos processos de tomada de decisão.

Desta forma, vêm tomando cada vez mais espaço as ferramentas de BI 2.0 que utilizam a Web 2.0 para buscar estas informações.

As redes sociais tornaram-se grandes centros de informações. Todos os dias, centenas de usuários e empresas vão à internet postar informações sobre diversos temas. Dentre estas informações estão gráficos, palavras, textos, áudios, vídeos e outros tipos

de informações que são usadas para expressar variados temas contemporâneos permitindo a compreensão de assuntos atuais de maneira rápida. Porém apenas existir estas informações, não é suficiente para uma análise, pois devido à grande quantidade de dados acabam prejudicando e desvirtuando os resultados. Com a finalidade de analisar estes outros tipos de dados, as ferramentas de BI 2.0 coletam estes dados, armazenam em estruturas de dados organizadamente e em seguida apresentam estes dados de forma simples aos usuários.

As ferramentas de BI 2.0 possibilitam a coleta de informações atuais de forma rápida, pois utilizam como fontes de dados a própria Web 2.0. Eventos cotidianos e fatos ocorrem e segundos depois estão sendo publicados pelos usuários, fazendo com que estas ferramentas possuem informações, muitas vezes, mais recentes do que as encontradas nas fontes estruturadas.

Com o objetivo de demonstrar um exemplo de sistema para suprir esta necessidade, este artigo tratou como assunto principal o desenvolvimento de uma solução de BI 2.0 para análise de dados provenientes de fontes não estruturadas, neste caso o Twitter. Com este exemplo de sistema e com os resultados obtidos, demonstrados no capítulo 5 deste trabalho, foi possível constatar a possibilidade de analisar assuntos atuais de forma rápida e fácil, permitindo que processos de tomada de decisão anteriormente dificultosos, sejam apoiados por um novo tipo de abordagem com a demonstração de dados obtidos pela opinião de diversas pessoas, cada quais com suas impressões sobre o tema, aproximando o analista dos usuários que fazem menção sobre o tema da análise.

Pelo fato do trabalho proposto apresentar algumas limitações, considerando-se que a única fonte de dados coletada é o Twitter, como possível proposta futura pode-se apontar a utilização de outras fontes de dados, por exemplo, o Facebook, com a finalidade de ampliar ainda mais as informações disponíveis para o usuário nos processos decisórios. Sugere-se ainda, em decorrência da necessidade do mercado, um módulo mobile, contemplando as principais funcionalidades do sistema, como pesquisa de informações, geração de gráficos e relatórios.

7.Referências

Bressan, R. T. (2007). Dilemas da rede: Web 2.0, conceitos, tecnologias e modificações. Intercom - Sociedade Brasileira de Estudos Interdisciplinares da Comunicação, Santos.

Ceci, F. (2012). Business Intelligence. 2. ed. Palhoça: Livro Digital. 176 p.

Ceci, F. (2010). Um modelo semiautomático para a construção e manutenção de ontologias a partir de bases de documentos não estruturados. Dissertação (Mestrado em Engenharia e Gestão do Conhecimento) – Universidade Federal de Santa Catarina, Florianópolis.

Ceci, F., Pietrobon, R., & Gonçalves, A. L. (2012). Turning Text into Research Networks: InformationRetrieval and Computational Ontologies in the Creation of Scientific Databases. Retrieved June 14, 2015, from http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0027499

Côrtes, S. (2002). BI, Data Warehouse e Data Mining - Como a Tecnologia aumenta a Inteligência do Negócio. PUC-Rio.

Cruz, T. (2008). *BPM e BPMS* - Business Process Management e Business Process Management Systems (2ª Ed). Rio de Janeiro: Brasport.

Gonçalves, A. L. (2006). Um modelo de descoberta de conhecimento baseado na correlação de elementos textuais e expansão vetorial aplicado à engenharia e gestão do conhecimento. Tese de Doutorado, Universidade Federal de Santa Catarina(UFSC), Florianópolis, Santa Catarina, Brasil.

Martins, D. B. (2008). Aplicação de técnicas de distribuição e paralelismo em ambientes de BI 2.0 para suporte à qualidade de dados. Tese de Mestrado, Universidade Federal do Estado do Rio de Janeiro (UNIRIO), Rio de Janeiro, Brasil.

Mooney, R. J., & Nahm, U. Y (2003). Text Mining with Information Extraction. [Paper]. Department of Computer Sciences, University of Texas, Austin, Texas, EUA.

Moura, A. (2012). Da Web 2.0 à Web 2.0 móvel: implicações e potencialidades na educação. Revista Limite, volume (nº 4), p81-104.

Nelson, G. S. (2010). Business Intelligence 2.0: Are we there yet? Proceedings of the Thot Wave Technologies, Chapel Hill, North Carolina, EUA.

Pintas, J. T., & Siqueira, S. W. M. (2011). O papel da semântica no Business Intelligence 2.0: Um exemplo no contexto de um programa de pós-graduação. Proceedings of the VII Simpósio Brasileiro de Sistemas de Informação. Salvador, Bahia, Brasil.

Turban, Efraim et al. (2009). Business Intelligence*:* Um enfoque gerencial para a inteligência do negócio. Porto Alegre: Bookman.