

**UNIVERSIDADE ESTADUAL DO PIAUÍ – UESPI**  
**CAMPUS PROF. ALEXANDRE ALVES DE OLIVEIRA**  
**CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

**DANIELE DE ARRUDA CALIXTO**

**EXTRAÇÃO E ANÁLISE DE DADOS NA APLICAÇÃO DE ESTRATÉGIAS DE  
MARKETING NO FACEBOOK DE UMA INSTITUIÇÃO SOCIAL**

**Biblioteca UESPI PHS**  
Registro Nº \_\_\_\_\_  
CDD 001.535  
CUTTER C1532  
V. \_\_\_\_\_ EX. 1  
Data 28 / 01 / 2018  
Visto. \_\_\_\_\_

**PARNAÍBA**

**2015**

**DANIELE DE ARRUDA CALIXTO**

**EXTRAÇÃO E ANÁLISE DE DADOS NA APLICAÇÃO DE ESTRATÉGIAS DE  
MARKETING NO FACEBOOK DE UMA INSTITUIÇÃO SOCIAL**

Monografia submetida ao Curso de Bacharelado em Ciência da Computação da Universidade Estadual do Piauí, como parte dos requisitos para obtenção do título de Bacharel em Ciência da Computação.

Orientador: Prof. M.Sc. Dario Brito Calçada

**PARNAÍBA**

**2015**

C153e

Calixto, Daniele de Arruda

Extração e análise de dados na aplicação de estratégias de marketing no Facebook de uma instituição social /

Daniele de Arruda Calixto. – Parnaíba: UESPI, 2015.

44 f. : il.

Orientador: MSc. Dario Brito Calçada

Monografia (Graduação em Ciência da Computação) - Universidade Estadual do Piauí, 2015.

1. Extração de Conhecimento 2. Marketing 3. Rede Social 4. Facebook  
5. Análise de Dados 6. Gephi 7. Netvizz I. Calçada, Dario Brito II.  
Universidade Estadual do Piauí III. Título

CDD 001.535

**DANIELE DE ARRUDA CALIXTO**

**EXTRAÇÃO E ANÁLISE DE DADOS NA APLICAÇÃO DE ESTRATÉGIAS DE  
MARKETING NO FACEBOOK DE UMA INSTITUIÇÃO SOCIAL**

Monografia apresentada ao Curso de Bacharelado em  
Ciência da Computação da Universidade Estadual do  
Piauí, como parte das exigências da disciplina de  
Estágio Supervisionado, requisito parcial para  
obtenção do título de Bacharel em Ciência da  
Computação.

Orientador: Prof. M.Sc. Dario Brito Calçada


Monografia Aprovada em: **24 de julho de 2015.**

Banca Examinadora:




---

Prof. M.Sc. Dario Brito Calçada  
UESPI/Parnaíba – Orientador



---

Prof. Dr. Sérgio Barros de Sousa  
UESPI/Parnaíba – Avaliador



---

Prof. Esp. Paulo César Coutinho dos Santos  
CEAD/UAB/UEPI/Teresina – Avaliador

Dedico esta monografia a minha família que foram os principais incentivadores durante esses quatro anos de curso. Aos meus amigos, que sempre me motivaram e incentivaram durante essa etapa.

## AGRADECIMENTO

À minha família, em especial aos meus pais pelo amor e dedicação concedida e pelos cuidados de sempre.

Ao meu orientador prof. Dario Calçada, pelo suporte no pouco tempo que lhe coube, por sua ajuda e incentivo neste trabalho.

A turma de Computação 2015.2, que fizeram esses quatro anos na universidade serem mais divertidos, tornando excelente esse tempo de convivência.

Aos meus amigos Diego e Thiago, por todo o apoio e incentivo de sempre, por não me deixarem ir para o lado negro da força, sempre me lembrando dos ensinamentos Jedi.

Aos amigos André e Cynthia, pelos bons conselhos e pelas conversas motivadoras nos períodos mais complicados.

A todos os amigos que estiveram presentes durante esses anos de curso e me incentivaram e apoiaram.

Aos professores que me incentivaram e inspiraram a concluir essa etapa.

Agradeço a todos que de forma direta ou indireta fizeram parte da minha formação.

**“A mente que se abre a uma nova ideia jamais voltará ao seu tamanho original”. (Albert Einstein)**

## RESUMO

A extração e exploração de dados são de grande relevância, principalmente quando utilizadas em análises de estratégias de marketing em redes sociais, pois possibilita descobrir padrões em publicações e usuários influentes na rede. Considerando que as redes sociais limitam as informações que possuem um grau de importância maior quando utilizadas para fins relacionados ao marketing, é por meio da exploração de dados utilizando planos de marketing que será viável obter um maior alcance de público na instituição. O presente trabalho mostra a extração e análise de conhecimento em uma página de uma instituição social no *Facebook* com a finalidade de analisar os dados antes e depois da aplicação de estratégias de marketing. No desenvolvimento dessa pesquisa foi extraído conhecimento por intermédio do aplicativo *Netvizz* incluso na rede social, onde os dados foram divididos em períodos de tempo diferente, antes e após a aplicação de técnicas de marketing. O software *Gephi* viabilizou a exploração do conhecimento, os dados foram comparados com a finalidade de identificar os cinco usuários mais influentes e as cinco publicações com maior alcance em cada período e durante o tempo de existência da página, sendo analisados também os resultados das estratégias de marketing propostas com base nos dados analisados de períodos anteriores. A técnica aplicada se mostrou eficaz, tornando-se uma forma de obter um conhecimento amplo sobre os usuários, as publicações e suas conexões no *Facebook* para assim aplicar técnicas de marketing visando propagar informações e alcançando um maior público.

**PALAVRAS CHAVES:** Extração de Conhecimento. Marketing. Rede Social. Facebook. Análise de Dados. Gephi. Netvizz.



## ABSTRACT

The data extraction and exploration are of great importance, particularly when used in analysis of marketing strategies in social networks, it allows discover patterns in influential publications and users on the network. Whereas social networks limit the information they possess a higher degree of importance when used for purposes related to marketing, is through data mining using marketing plans that will be feasible to obtain a greater of audience reach in the institution. This work shows the knowledge extraction and analysis on a page of a social institution in *Facebook*, in order to analyze the data before and after the application of marketing strategies. In the development of this research the knowledge extracted through the application *Netvizz* included in the social network, where the data were divided into different periods of time before and after application of marketing techniques. The *Gephi* software enabled the exploration of knowledge , data were compared in order to identify the five most influential users and the five publications with greater reach in each period and during the life time of the page, also were analyzed the results of strategies marketing proposals according with the analyzed data of previous periods. The applied technique if showed effective, making it a way to get a broad knowledge of users, publications and connections for *Facebook* to so apply marketing techniques aimed at spreading information and reaching a wider audience.

**KEY-WORDS:** Knowledge Extraction. Marketing. Social Network. Facebook. Data analysis. Gephi. Netvizz.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Etapas do processo de extração do conhecimento	18
Figura 2 – Grafo gerado em estado bruto	21
Figura 3 – Grafos após a aplicação de algoritmos de layout	22
Figura 4 – Visão geral do trabalho desenvolvido	23
Figura 5 – Dados em estado bruto	25
Figura 6 – Interface Gephi	26
Figura 7 – Gráfico: Publicações em períodos de tempo diferentes	29
Figura 8 – Grafo de divisão dos nós da rede no período de 2014.1 ( <i>Gephi</i> )	30
Figura 9 – Gráfico: Envolvimento dos usuários	31
Figura 10 – Gráfico: Publicações do segundo semestre de 2014 e maio de 2015	32
Figura 11 – Gráfico: Comparação das publicações nos meses de maio e junho	33
Figura 12 – Grau de entrada ( <i>Gephi</i> )	34
Figura 13 – Gráfico: Publicações com maior alcance	35
Figura 14 – Gráfico: Usuários mais influentes	35
Figura 15: Modularidade ( <i>Gephi</i> )	36
Figura 16: Tipos de nós ( <i>Gephi</i> )	37

## **LISTA DE TABELAS**

Tabela 1 – Períodos de coleta de dados e quantidade de nós e arestas

25

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	11
<b>2 ANÁLISE DE REDES SOCIAIS E MARKETING</b> .....	14
2.1 MARKETING .....	14
2.2 FACEBOOK.....	15
<b>3 EXTRAÇÃO, VISUALIZAÇÃO E ANÁLISE DE DADOS</b> .....	17
3.1 EXTRAÇÃO DE CONHECIMENTO .....	17
3.2 VISUALIZAÇÃO DE DADOS .....	19
3.3. NETVIZZ .....	20
3.4 GEPHI .....	20
<b>4 MATERIAIS E MÉTODOS</b> .....	23
4.1 EXTRAÇÃO DE DADOS NO FACEBOOK .....	24
4.2 ANÁLISE VISUAL DOS DADOS .....	26
<b>5 RESULTADOS E DISCUSSÃO</b> .....	29
<b>6 CONSIDERAÇÕES FINAIS</b> .....	38
<b>REFERÊNCIAS</b> .....	40

## 1 INTRODUÇÃO

Uma rede social retrata uma estrutura composta por pessoas ou organizações, conectadas por um ou vários tipos de relações, que partilham valores e propósitos comuns (FERREIRA, 2011). De acordo com Miranda e colaboradores (2009), as redes sociais viabilizam a existência de um poder de influência amplo entre as pessoas envolvidas de acordo com o contato direto que possuem. Essa influência cresce ou diminui por meio da utilização dos mecanismos oferecidos pela rede social, que podem ser mensagens, grupos, e outras publicações de interesses pessoais.

O Marketing Viral atua no sentido estratégico de capturar a atenção, possibilitando um receptor a se tornar um emissor, buscando influenciar sua rede de convivência social online. (ANDRADE; MAZZON; KATZ, 2006). Desta forma utilizando táticas de marketing em conjunto com os mecanismos das redes sociais é possível obter um maior alcance da informação.

Segundo Panteli (2009), a rede social *Facebook* possui uma participação massiva, devido a possuir um vasto número de usuários registrados, quando essa quantidade de usuários é elevada causa um aglomerado, formando grupos que tendem a agir de forma conjunta.

Para compreender melhor o efeito que a divulgação de um empreendimento pode proporcionar em redes sociais é necessário conhecer o público alvo e utilizar estratégias de marketing, com a finalidade de alcançar uma maior disseminação do projeto, visando assim conquistar mais apoiadores. Sendo fundamental descobrir o perfil de usuários que interagem com maior frequência, os dados que obtém maior destaque e como isso ocorre, o crescimento do projeto na mídia de acordo com a interação dos usuários de redes sociais e ter conhecimento da influência dos dados durante o período de propagação de informações (TOMÁÉ; ALCARÁ; CHIARA, 2005).

De acordo com Benevuto e colaboradores (2011) são realizadas pesquisas online, como enquetes, questionários, para avaliar a interação do público com determinado projeto, entretanto estas pesquisas online não são suficientes para uma compreensão eficaz dos dados obtidos por redes sociais.

A partir da problemática no qual se busca conseguir um entendimento melhor sobre as informações coletadas no *Facebook* e assim analisar as mesmas de forma mais precisa, a exploração e análise dos dados de redes sociais por intermédio da coleta e utilização de um software específico para visualização de dados apresentam-se como uma alternativa

para ter um controle maior sobre as informações. A coleta dos dados na rede social *Facebook* possibilita ter acesso a diferentes dados de determinado período de tempo, visto que estes só podem ser extraídos a partir de uma ferramenta específica. (ROSA, et al., 2012).

A exploração e análise de dados proporcionam: a descoberta de padrões e a definição dos mesmos, além de permitir à filtragem de dados específicos. A realização do agrupamento de informações possibilita a visualização de dados de uma forma ampla, sendo isto de grande importância, pois proporciona definir o perfil do público, delimitar estratégias de marketing na mídia e acompanhar o crescimento da campanha de divulgação realizada nas redes sociais (FREITAS, et al., 2008).

A principal finalidade deste trabalho é mostrar a importância da utilização de exploração e análise de dados em uma rede social sobre determinada instituição. Para atender ao objetivo geral, alguns objetivos específicos foram propostos, como: definir os usuários mais influentes; analisar e criar estratégias de marketing a partir dos dados coletados; analisar o crescimento do projeto de acordo com a interação dos usuários nas redes sociais; e visualizar em grafo os dados nas redes sociais relacionados ao projeto.

A seleção do *Facebook* para a coleta de dados ocorreu devido ao fato do empreendimento sem fins lucrativos ser uma Obra Social que possui conta e interação com usuários dessa rede, além de possuir um vasto número de usuários e ser classificada entre as redes sociais mais acessadas no mundo.

Para a coleta de dados no *Facebook* das Obras Sociais Luz da Esperança do município de Parnaíba no estado do Piauí, utilizou-se uma ferramenta chamada *Netvizz*, a qual é uma aplicação que está inclusa na própria rede social e possui autorização para realizar coleta de dados em grupos ou páginas e a interação dos usuários com publicações, possibilitando determinar o período de coleta. Após a coleta de dados utilizou-se uma ferramenta chamada *Gephi* para a exploração e análise dos dados coletados. Essa ferramenta viabiliza a análise de extensas redes de dados, possui um motor de renderização 3D, gerando uma melhor visualização dos dados coletados em grafos, além de mostrar resultados estatísticos, permite a definição de padrões e o uso de algoritmos.

Esta monografia está dividida metodologicamente em seis capítulos, onde serão brevemente comentados.

O atual capítulo, “INTRODUÇÃO”, apresenta uma visão total dos principais pontos desta pesquisa, considerando alguns pontos importantes como a justificativa, problemática, objetivos e metodologia utilizada na resolução do problema.

O capítulo 2, “ANÁLISE DE REDES SOCIAIS E MARKETING”, mostra alguns conceitos relevantes para a compreensão do tema da atual pesquisa.

O capítulo 3, “EXTRAÇÃO, VISUALIZAÇÃO E ANÁLISE DE DADOS”, apresenta alguns conceitos importantes e ferramentas utilizadas neste trabalho, como o Software *Gephi*, mostrando características técnicas e alguns conceitos simples.

O capítulo 4, “MATERIAIS E MÉTODOS”, mostra como foi usada às ferramentas utilizadas neste trabalho, assim como detalhes técnicos. Neste capítulo será apresentado como os dados foram coletados, explorados e analisados.

O capítulo 5, “RESULTADOS”, mostra o desenvolvimento final deste trabalho, como tópicos sobre a identificação do público mais influente, publicações mais relevantes segundo a exploração e análise dos dados, bem como o resultado das estratégias de marketing conforme os dados coletados.

O último capítulo, “CONSIDERAÇÕES FINAIS”, contém algumas observações acerca do trabalho desenvolvido. São apresentadas também, algumas sugestões de melhorias para a exploração de dados em redes sociais no tópico de trabalhos futuros.

## 2. ANÁLISE DE REDES SOCIAIS E MARKETING

A análise de redes sociais (ARS) busca dentro de um determinado sistema identificar conexões (laços ou relações) e pontos (nós ou atores), e a partir disso, representar padrões entre relações que podem ser constantes ou imprevisíveis. (FAZITO, 2002)

As redes sociais viabilizam um ambiente comum de interesses, necessidades e objetivos semelhantes para a partilha de conhecimento, a interação e comunicação entre seus usuários (PETTENATI et al., 2006). Desta forma a comunicação oferecida por redes sociais possibilita a formação de conjuntos de usuários com algum interesse em comum, aproximando pessoas com o mesmo objetivo.

Redes sociais são representadas por grafos onde os nodos são os atores (geralmente pessoas) e as arestas são os relacionamentos entre esses atores. Podem apresentar desde conexões esparsas, como em árvores genealógicas, até conexões muito densas, como em redes de contato na internet. (FREITAS, et al., 2008)

Uma rede é caracterizada por um agrupamento de nós ligados por arestas. Dessa forma uma rede social é um conjunto de atores (pessoas, instituições ou grupos) e suas conexões (WASSERMAN; FAUST, 1994).

As mídias sociais contribuem para que os indivíduos se organizem de acordo com seus interesses, permitindo uma comunicação mais ampla em uma escala global. (CASTELLS, 2004). As plataformas de redes sociais modificaram a forma como os seus usuários interagem, ampliando a comunicação online e interações entre usuários, sendo que milhões de usuários compartilham informações e suas opiniões sobre diversos assuntos. (VAKALI, 2012).

### 2.1 MARKETING

Segundo Kotler (2000), o marketing é um meio social em que pessoas e grupos conseguem aquilo que precisam e o que desejam por meio da criação, oferta e negociação de serviços e produtos. Rosen (2001) afirma que cada cliente que recebe informações de um amigo pode distribuí-las para vários outros amigos em suas redes sociais, esta forma de marketing por meio da internet é chamada de marketing viral.

O marketing online atinge um grupo determinado de usuários, no qual consegue desenvolver um diálogo contínuo com eles, simplificando assim a comunicação sobre qualquer assunto relacionado à empresa (CINTRA, 2010).



O uso das estratégias digitais em campanhas de comunicação e marketing das empresas tem se diversificado pelo uso sistemático de ferramentas, como a integração da mobilidade e portabilidade de acesso à web e com a otimização dos sistemas de busca, instituindo novos pilares de comunicação integrada (OKADA e SOUZA, 2011).

O marketing digital utiliza as tecnologias como ferramentas principais para melhorar sua comunicação com clientes, os profissionais devem exercer o papel de disseminador da informação e utilizá-lo de forma coerente, adequando-se a mudanças quando necessário e visando sempre atender as necessidades dos clientes (LAS CASAS, 2006).

O marketing viral é um método de divulgação de determinada mensagem publicitária, com base em seu poder de alcance por meio de pessoas próximas, a divulgação em larga escala de uma mensagem ocorre devido a uma ação participativa de usuários em redes sociais (BARICHELLO e OLIVEIRA, 2006).

De acordo com Andrade e colaboradores (2006) o marketing viral atrai a atenção de consumidores, com o propósito de fazer do consumidor um operador de venda, buscando influenciar seus contatos em redes sociais. Conforme Wang e colaboradores (2008) com o compartilhamento de informações com amigos na internet aumenta o crescimento de benefícios estratégicos dentro de uma organização.

O marketing viral atua no sentido estratégico de capturar a atenção do consumidor e de fazê-lo, ao mesmo tempo, um agente de venda, ou dito de outra forma, possibilita a um receptor (passivo) a se tornar num emissor (ativo), buscando influenciar sua rede de convivência social online (ANDRADE; MAZZON e KATZ, 2006).

Conforme Barichello e Oliveira (2006) o marketing viral é uma das formas em que os consumidores se comunicam entre si a existência de certa marca. Ao oferecer aos usuários o poder de comunicação e propagação de alguma mensagem, a empresa consegue com que seu anúncio alcance um grande número de pessoas a um custo barato e a uma velocidade mais rápida do que no uso da mídia tradicional.

## 2.2 FACEBOOK

O *Facebook* foi criado pelo americano Mark Zuckerberg, lançado em 2004, sendo umas das redes sociais com maior número de usuários do mundo. Funciona por meio de perfis de usuários e comunidades (páginas e grupos), onde em cada perfil é possível acrescentar aplicativos existentes ou o próprio usuário pode criá-los para o site (RECUERO, 2009).

## Segundo informações disponíveis na rede social *Facebook* :

Pessoas de todas as idades e de todas as partes do mundo acessam o Facebook para entrar em contato com amigos, familiares e coisas importantes para elas. Mais de 1,35 bilhão de pessoas visitam o Facebook todos os meses, e 64% delas acessam o Facebook diariamente (FACEBOOK, 2015).

A comunicação no *Facebook* não é realizada apenas por meio de perfis pessoais, a rede social permite a criação de grupos e páginas com foco em determinado assunto. Segundo o *Facebook* (2015), as páginas são semelhantes aos perfis pessoais, porém possuem ferramentas exclusivas que ajudam na conexão de usuários com algum tópico importante para ele, seja sobre cultura, alguma empresa ou organização, entre outros.

As Páginas do Facebook são gratuitas e fáceis de configurar. Pode-se usar uma Página para alcançar objetivos específicos. Por exemplo, pode-se publicar imagens de produtos mais recentes para ajudar a impulsionar as vendas. Ou compartilhar um vídeo pequeno para ajudar as pessoas a se lembrar de uma empresa. Quando as pessoas curtem uma Página ou as suas publicações, os amigos delas também podem descobrir este mesmo perfil no Facebook (FACEBOOK, 2015).

De acordo com o *Facebook* (2015), as páginas tornam uma empresa mais:

- Conectada, possibilitando a conversa com colaboradores que curtem a página e visualizam as publicações;
- Localizável, quando algum cliente pesquisar pelo nome da empresa encontrará informações como serviços, endereço, entre outros dados;
- Oportuna, por meio da página é possível atrair um maior número de pessoas, frequentemente, com publicações personalizadas conforme as necessidades e interesse da empresa;

### 3 EXTRAÇÃO, VISUALIZAÇÃO E ANÁLISE DE DADOS

O sucesso das redes sociais fez com que a análise dessas redes ganhasse uma atenção maior por parte de organizações, visto que é possível realizar uma análise para melhorar suas estratégias de marketing. (LEITE et al., 2013). O grande número de usuários proporcionou um amplo acervo de informações sobre atividades realizadas nessas plataformas, gerando assim muitas informações. Segundo Rezende (2005), as informações adquiridas em redes sociais ajudam na análise de padrões, no qual contribuem para obter um maior conhecimento sobre determinado tema.

As mídias sociais utilizam-se de Interface de Programação de Aplicação (APIs) específicas, que oferecem um conjunto de rotinas e padrões estabelecidos por um software para a utilização das suas funcionalidades por aplicativos que não pretendem envolver-se em detalhes da implementação do software, mas apenas usar seus serviços. Para ter acesso ao banco de dados das mídias sociais, é necessário que seja criado um aplicativo no site da mídia, normalmente localizado na parte de desenvolvedor onde o site libera um *token* de acesso (POLONI, TOMAÉL, 2014).

Segundo Poloni e Tomáel (2014), a API do *Facebook* possui limitações, pois permite acesso somente a informações de amigos ou pessoas que autorizem que um aplicativo faça uso de suas informações.

#### 3.1 EXTRAÇÃO DE CONHECIMENTO

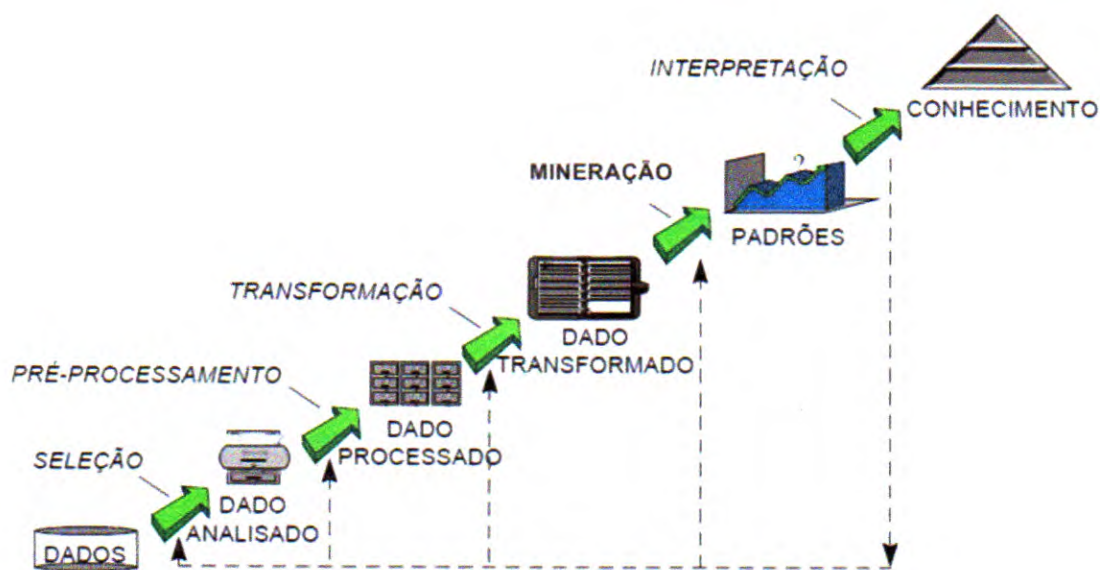
A extração de conhecimento utiliza uma base de dados em estado bruto, estes dados passam por um conjunto de etapas com o objetivo de alcançar um conhecimento sobre determinado assunto (FAYYAD, 2002).

O conhecimento descoberto deve apresentar três atributos: deve ser eficiente, aplicável a diferentes tipos de dado e ajustável. O processo de descoberta de conhecimento deve atender a três critérios: deve ser preciso, compreensível por usuários, além de ser útil e interessante (FREITAS, 2000).

Na Figura 1 é possível ver as etapas do processo de extração de conhecimento conforme descrito abaixo por Almeida e colaboradores (2013), as etapas são:

- Seleção, nessa etapa é selecionada um conjunto de dados para que possa ser analisado;
- Pré-processamento, os dados são organizados e tratados;
- Transformação, os dados são armazenados e formatados adequadamente;

- Mineração, nesta etapa é utilizada técnicas inteligentes;
- Interpretação, o conhecimento adquirido é analisado.



**Figura 1:** Etapas do processo de extração do conhecimento

Fonte: (FAYYAD, 2002)

De acordo com Kremer (1999):

Prospecção de conhecimento em bases de dados (*Knowledge Discovery in Databases - KDD*) é um processo que envolve a automação da identificação e do reconhecimento de padrões em um banco de dados (...). Sua principal característica é a extração não trivial de informações a partir de uma base de dados de grande porte. Essas informações são necessariamente implícitas, previamente desconhecidas, e potencialmente úteis.

O *Knowledge Discovery in Databases* (KDD) busca alcançar conhecimento a partir de uma grande massa de dados, esse processo envolve as áreas de visualização de dados, análise de base de dados, aprendizado de máquina, sistemas especialistas, além dos dados de estatística. (AZEVEDO, 2011).

A mineração de dados (*Data Mining*) é uma etapa do processo de KDD que se associa a análise de dados e a utilização de ferramentas com o objetivo de encontrar padrões (características e regras) em um grande conjunto de dados (DINIZ, LOUZADA-NETO, 2000). *Data Mining* é basicamente um processo de extração de informações desconhecidas e úteis, a partir da aplicação de técnicas específicas em uma base de dados, com o propósito de

descobrir relações, padrões entre os dados para assim ter controle ou melhorar algum processo (MARTINS et al., 1999).

A mineração de dados possui técnicas que são aplicadas na descoberta de conhecimento em banco de dados para extrair determinadas informações escondidas em grandes bases de dados, por meio da determinação de padrões, associações e classificações (GOEBEL; GRUENWALD, 1999). A mineração de dados é uma das etapas do KDD que converte dados brutos em conhecimento, possui técnicas preditivas que em tarefas de classificação buscam construir modelos que possam apresentar uma melhor combinação de relacionamento entre um conjunto de dados, chamado de atributo explanatório e um dado de interesse, chamado de atributo preditivo (TAN, 2006).

### 3.2 VISUALIZAÇÃO DE DADOS

A visualização é uma técnica que apresenta múltiplos aspectos de informações reunidas, sendo necessária a utilização de algoritmos especificados para demonstrar eficientemente os dados (BARTOLOMEU, 2002). A aplicação de técnicas de visualização de dados pode ocorrer em qualquer etapa no processo de descoberta de conhecimento, essas técnicas proporcionam a representação mais exata sobre as informações e conhecimentos extraídos (YAMAGUCHI, 2010).

Segundo Freitas e colaboradores (2008):

Visualização de redes sociais é uma subárea de visualização de informações, onde o conjunto de dados a ser visualizado e explorado é um grafo, com os nodos representando entidades sociais (pessoas, instituições, grupos, etc) e arestas representando os relacionamentos existentes. Para a visualização, portanto, o grafo deve ter sido previamente extraído da base de dados da aplicação.

De acordo com Bruzzese e Buono (2004) pelo uso de grafos é possível visualizar as regras de associação por meio de diferentes aspectos. Hinnerburg e colaboradores (1999) afirma que é necessário combinar técnicas de visualização com algoritmos de agrupamento aplicados em base de dados. Para Backer (1995), a classificação ou agrupamento de objetos em grupos é bastante comum e necessária devido ao amplo número de informações disponíveis.

Uma rede é um grafo sendo este um conjunto de vértices (ou nós) e um conjunto de arestas (ou arcos) que conectam esses vértices, as arestas possibilitam que exista algum tipo de relação entre dois vértices de acordo com as características da rede. O grafo pode ser

direcionado ou não, um grafo direcionado também chamado de dígrafo, cada aresta tem um sentido que conecta um vértice origem a um vértice destino. Um vértice qualquer possui um grau no qual é determinado de acordo com o número de arestas que conectam aquele vértice, a distribuição de graus é uma função de distribuição probabilística no qual indica as chances de um determinado vértice possuir um grau fixo (METZ, et. al., 2007).

A divisão de uma grande quantidade de objetos em grupos que possuem semelhanças entre seus objetos é um método chamado análise de agrupamento (*clusterização*) (SOUKUP; DAVIDSON, 2002). Na análise de *cluster* é verificado se existe diferentes grupos em um conjunto de dados, caso existam, são definidos quais são esses grupos. (SFERRA e CORRÊA, 2003)

### 3.3 NETVIZZ

O *Netvizz* é uma ferramenta que possibilita a extração de dados de diferentes seções do *Facebook* (grupos, páginas) com a finalidade de análise, são criados arquivos em formato gdf (formato de texto que especifica um gráfico), esses arquivos gerados podem ser analisados e visualizados usando um software de visualização de gráfico. (NETVIZZ, 2015)

Segundo Rosa e colaboradores (2012), o *Netvizz* é um aplicativo integrado na rede social *Facebook*, no qual gera um arquivo de alguma seção da rede social incluindo os “pesos” das relações para gerar grafos. O aplicativo ainda gera um arquivo específico para a visualização de grafos. De acordo com Santos (2014), o *Netvizz* possui um mecanismo que extrai dados de grupos e páginas, possibilitando a mineração da relação entre páginas e verificando o envolvimento de usuários com postagens.

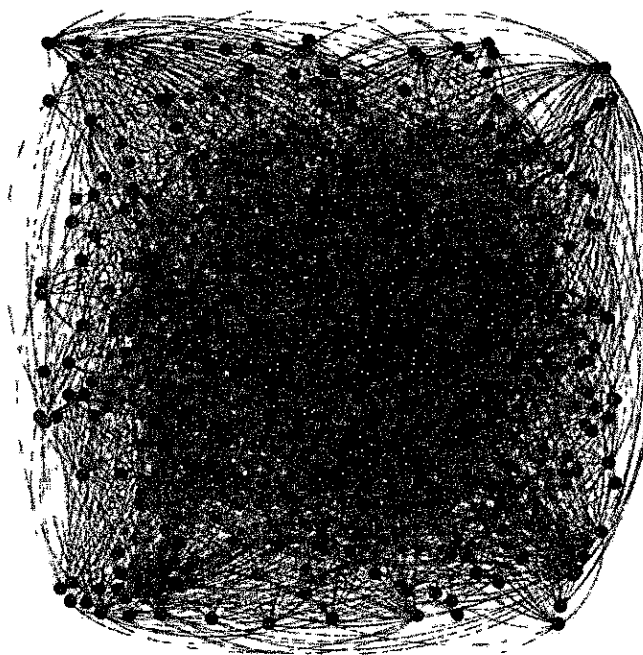
Os dados podem ser extraídos conforme a participação do usuário em páginas ou grupos, ou seja, se o usuário não está presente em um grupo ou participa de uma página não pode ter acesso aos dados destes locais, o arquivo gerado pode ser explorado e analisado com o uso de uma ferramenta específica (RIEDER, 2013).

### 3.4 GEPHI

*Gephi* é uma ferramenta para exploração de redes complexas, análise exploratória de dados e visualização, sendo complementar para as estatísticas tradicionais e tendo como objetivo ajudar a encontrar padrões, isolar singularidades ou falhas durante a organização dos dados (GEPHI, 2015). O software *Gephi* de código aberto é usado para a análise de redes, possibilitando o acesso amplo aos dados da rede, a filtragem, visualização, agrupamento de

dados e permitindo trabalhar com conjunto de dados complexos por meio de seus algoritmos, assim produzindo excelentes resultados visuais. Com a ferramenta *Gephi* é gerado um grafo hierárquico, no qual é possível aplicar vários algoritmos para calcular quais os dados mais influentes da análise realizada. (BASTIAN; HEYMANN; JACOMY, 2009)

De acordo com Rosa e colaboradores (2012) o software *Gephi* permite interação do usuário com a representação gerada por meio dos dados, possibilitando a manipulação da estrutura para revelar características por vezes não descobertas nos dados brutos. Utilizando essa ferramenta é possível transformar um arquivo criado no *Netvizz* a partir dos dados extraídos no *Facebook* em um grafo, no qual demonstra uma rede complexa de dados que a princípio é gerado em estado bruto, sem formatação. A figura 2 mostra um grafo em estado bruto elaborado pelo software *Gephi* por meio de dados de redes sociais.



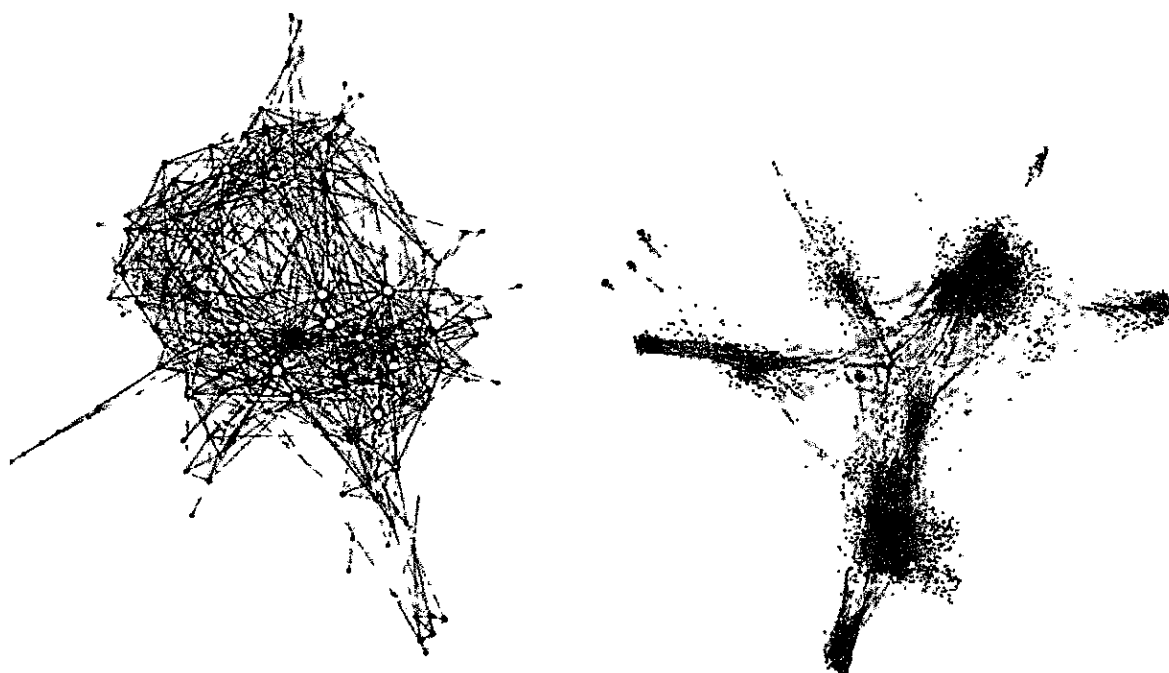
**Figura 2:** Grafo gerado em estado bruto.

Fonte: (ROSA, 2012).

A ferramenta *Gephi* possui algoritmos desenvolvidos por seus criadores, como por exemplo, o *Force Atlas*, que consegue organizar de forma rápida grupos de indivíduos. Os algoritmos possuem várias especialidades com o objetivo de organizar os dados para obter uma melhor análise e visualização (BASTIAN; HEYMANN and JACOMY; 2009).

De acordo com *Gephi* (2015), este software possui um rápido motor de visualização que acelera a descoberta de padrões em grandes gráficos, dispondo de ferramentas avançadas para a manipulação de dados e gráficos, algumas delas:

- Algoritmos de layout: contribuem na eficiência e qualidade no agrupamento e visualização. A figura 3 mostra grafos após a aplicação de algoritmos de layout.



**Figura 3:** Grafos após a aplicação de algoritmos de layout.

Fonte: (GEPHI, 2015)

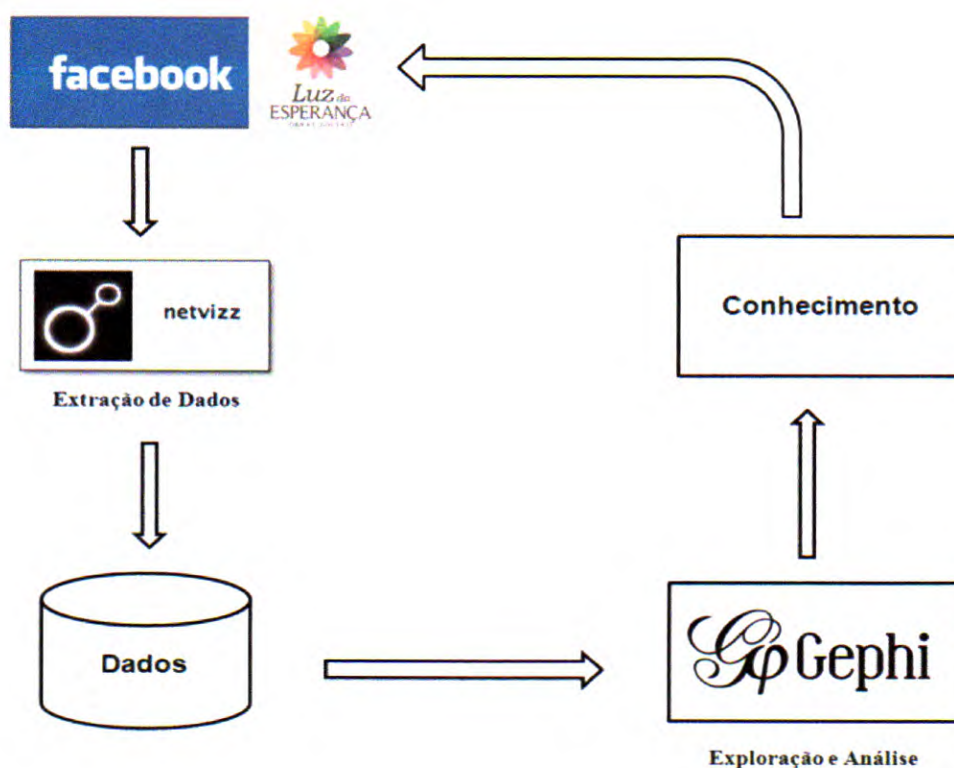
- Métrica: possui um quadro de estatísticas que fornecem métricas mais comuns para a análise de redes sociais (ARS) e redes de escala livre, possibilitando utilizar a proximidade, diâmetro, coeficiente de *clustering*, detecção de comunidades, entre outros;
- Análise de Rede Dinâmica: estruturas de redes sociais podem ser filtradas com a linha do tempo;
- Filtragem dinâmica: a rede é filtrada para selecionar os nós e arestas com base na rede de dados.
- Cartografia: utiliza dados de classificação e partição para representar a rede, personaliza cores, tamanho ou rótulos para demonstrar o funcionamento da rede;



## 4 MATERIAIS E MÉTODOS

Neste capítulo será detalhado como foi realizada a extração e análise de conhecimento de uma rede social e como esses dados coletados podem contribuir no crescimento da página com o uso do marketing.

O foco deste trabalho é voltado à análise de dados da página do *Facebook* das Obras Sociais Luz da Esperança, localizada no município de Parnaíba no estado do Piauí, desta forma, ocorreu à coleta de dados a partir de um aplicativo que está incluso na própria rede social, no qual possibilitou a extração do conhecimento de determinada página. A figura 4 mostra um diagrama que dá uma visão ampla do trabalho desenvolvido.



**Figura 4:** Visão geral do trabalho desenvolvido.

Fonte: Próprio Autor (2015).

Esses dados foram extraídos por meio da ferramenta *Netvizz* em determinados períodos de tempo. Os dados foram explorados no software *Gephi*, os quais foram comparados e analisados com a finalidade de identificar conhecimento sobre o público e o desempenho ocorrido por intermédio das estratégias de marketing usadas ao longo do tempo.

Para uma melhor compreensão das técnicas aplicadas e da análise dos dados, o trabalho foi dividido em **Extração de Dados no Facebook**, o qual equivale os requisitos utilizados na extração de conhecimento de uma página e a **Análise Visual dos Dados**, no qual o conhecimento adquirido foi explorado e usado no plano de marketing da instituição.

#### 4.1 EXTRAÇÃO DE DADOS NO FACEBOOK

Na extração de dados foi selecionado um aplicativo que possibilitou extrair mais dados importantes em comparação a outras ferramentas, desta forma utilizou-se o *Netvizz* na versão 1.2 projetado para funcionar com a versão 2.2 da *Application Programming Interface* (API) do *Facebook*, no qual é um aplicativo contido dentro da rede social necessitando de permissão do usuário para ser autorizado o uso. Após a permissão do usuário, o mesmo é direcionado a uma página que mostra informações sobre o aplicativo e opções para a extração dos dados, permitindo realizar a coleta em grupos, páginas e sobre determinado tema pesquisado.

As funções desse aplicativo são de acordo com o conteúdo que a API do *Facebook* permite extrair, não sendo possível conseguir o nome dos usuários, onde os mesmos são identificados apenas com um ID de acordo com o seu perfil apenas com o uso do aplicativo, é fornecido um ID para cada usuário de acordo com seu perfil. Todos os dados analisados foram coletados a partir de um processo de extração de conhecimento, nesse procedimento foram selecionadas as principais formas de interação entre usuários e a página, ou seja, publicações, comentários, curtidas e compartilhamentos.

O conhecimento analisado neste trabalho foi extraído da página Obras Sociais Luz da Esperança na rede social *Facebook*, a extração dos dados foi viável por meio do aplicativo *Netvizz*, o qual gerou um arquivo em formato *.gdf* contendo os dados em tabelas e grafos, sendo estes dados explorados no software *Gephi*, cada nó representou uma publicação ou um usuário, as arestas representaram as interações dos usuários com as postagens da página.

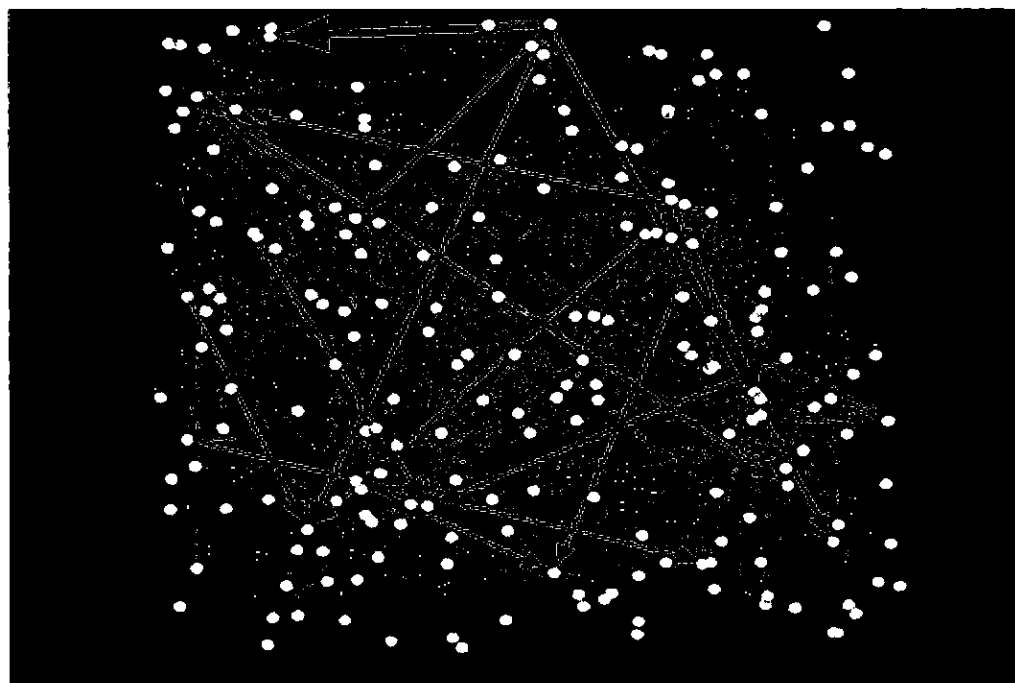
A Tabela 1 mostra os períodos de extração de dados e a quantidade de nós e arestas coletados em cada espaço de tempo.

Período de Tempo	Nós	Arestas
2º Semestre – 2013	80	120
1º Semestre – 2014	212	582
2º Semestre – 2014	165	438
01/01 até 30/04/2015	241	388
Maio	130	387
01/06 até 15/06/2015	105	173
22/08/2013 até 15/06/2015	682	2089

**Tabela 1** – Períodos de coleta de dados e quantidade de nós e arestas.

Fonte: Próprio Autor (2015).

Após a extração o aplicativo gerou um arquivo em formato *.gdf*, este arquivo mostra em tabelas e grafos os dados. A Figura 5 mostra os dados em estado bruto.



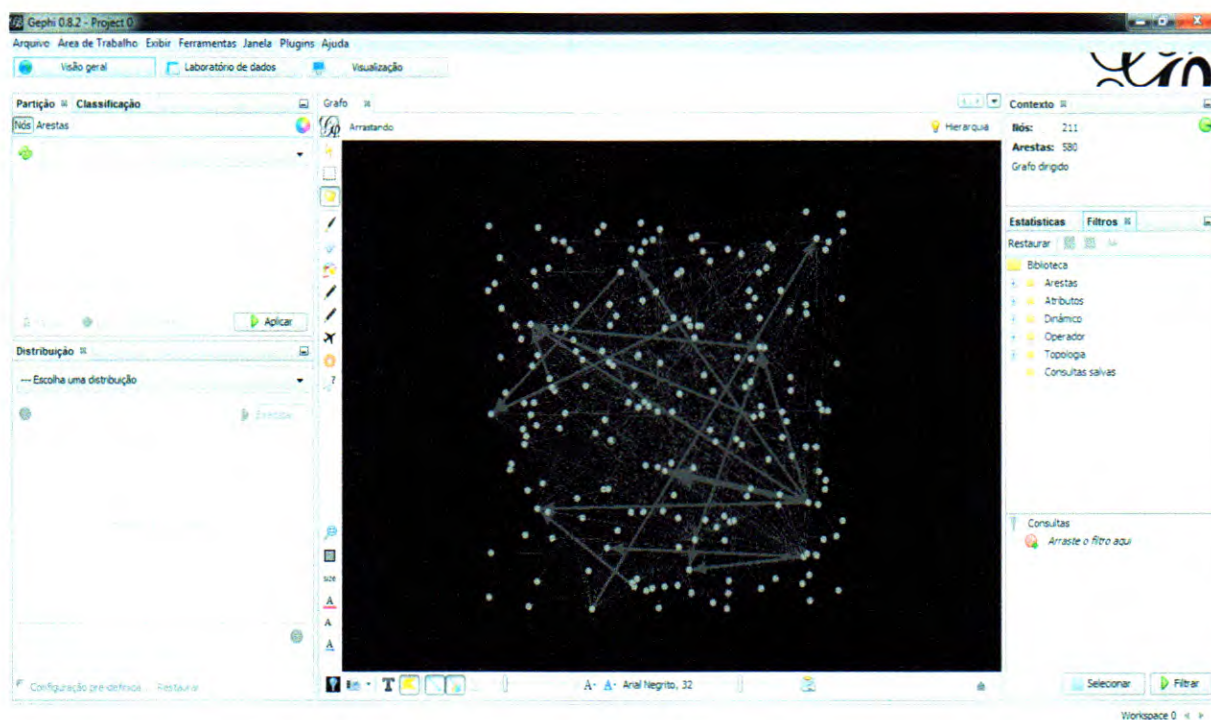
**Figura 1:** Dados em estado bruto.

Fonte: Próprio Autor (2015).

## 4.2 ANÁLISE VISUAL DOS DADOS

O arquivo gerado pelo aplicativo *Netvizz* foi explorado na versão 0.8.2 do software *Gephi*, assim sendo possível realizar a análise dos dados possibilitando obter diversas informações importantes para o crescimento da página na rede social. Esta ferramenta possibilitou a utilização de várias técnicas para identificar as conexões dentro de uma rede de dados, assim viabilizando a definir o público, mostrar a influência de usuários, comparando períodos diferentes da página, determinando e avaliando estratégias de marketing, bem como permitindo ampliar o número de usuários que apoiam a instituição.

A ferramenta *Gephi* possibilita ter uma visão geral do grafo aplicando técnicas para organizá-lo, extraíndo sentido da rede formando grupos que se conectam entre si. No laboratório de dados é possível verificar todas as informações sobre cada aresta e nó, sendo permitido definir padrões para a visualização dos dados e editar informações como o tamanho e cor de um nó. A visualização possibilita representar a rede de uma forma que se torne mais compreensível, possibilitando mostrar ou não o nome de cada nó e como a informação vai aparecer no grafo. A Figura 6 mostra a interface do software *Gephi*.



**Figura 2:** Interface Gephi.  
Fonte: Próprio Autor (2015).



O software *Gephi* possibilita compreender os dados extraídos e as informações mais relevantes sobre usuários e publicações. Foram definidas técnicas da ferramenta que são importantes na análise dos dados para alcançar os objetivos estabelecidos, logo foram utilizadas as estatísticas de grau médio, grau médio ponderado, grau de entrada, grau de saída, centralidade de autovetor, modularidade, e alguns algoritmos de layout.

O grau de entrada mostrou o número de conexões que um nó recebe de outro, no grau de saída foi possível verificar a conexão entre os nós por meio do número de arestas que ele envia para outros nós. A modularidade demonstrou o número de componentes agrupados na rede analisada, o algoritmo de layout *Force Atlas* tem como característica a velocidade, a gravidade e a repulsa, evitando sobreposição de nós e ajustes de tamanho, tornando o grafo mais compreensível. O algoritmo de layout Fruchterman-Reingold distribuiu os vértices de forma igualitária, diminuiu o cruzamento das arestas, fazendo um com mais simetria e contribuindo na compreensão do grafo.

HITS (Hyperlink-Induced Topic Search) é um algoritmo que determina os valores para cada nó classificando em: Hubs, no qual o valor das arestas que saem de um nó em direção aos outros, sendo assim quanto mais um nó conecta diferentes partes da rede, maior será seu valor como conector nessa rede; Autoridade, um nó que se conecta com mais hubs tem mais autoridade.

A utilização desses métodos citados proporcionou a interpretação do conhecimento adquirido pela exploração e análise dos dados. Dentre as funções e algoritmos utilizados, foram também estabelecido parâmetros para a análise da rede, entre eles a identificação de usuários e publicações, comentários, curtidas e compartilhamento.

Foram determinados os cinco usuários e cinco publicações mais importantes em cada período desde a criação da página até o mês de junho de 2015, onde os usuários foram classificados conforme o seu grau de envolvimento e as publicações de acordo com o alcance, sendo determinadas também informações relativas à data da postagem na página. Estes dados foram analisados e selecionados para assim demonstrar a influência de cada indivíduo na página. Foi minerado o conteúdo com a finalidade de descobrir informações relevantes para o uso em estratégias de marketing com a finalidade de obter a divulgação e o crescimento da página no *Facebook*.

Visto que a página Obras Sociais Luz da Esperança pertence a uma instituição sem fins lucrativos e que precisa de colaboradores para realizar seu trabalho social, após a análise do conhecimento, realizou-se a aplicação de estratégias de marketing com o objetivo de alcançar um maior público na rede social. Após o tempo em que essas estratégias foram

aplicadas, ocorreu à extração de dados desse período e análise determinando os principais usuários e publicações, em seguida ocorreu à comparação dos dados antes e depois da aplicação de técnicas de marketing.

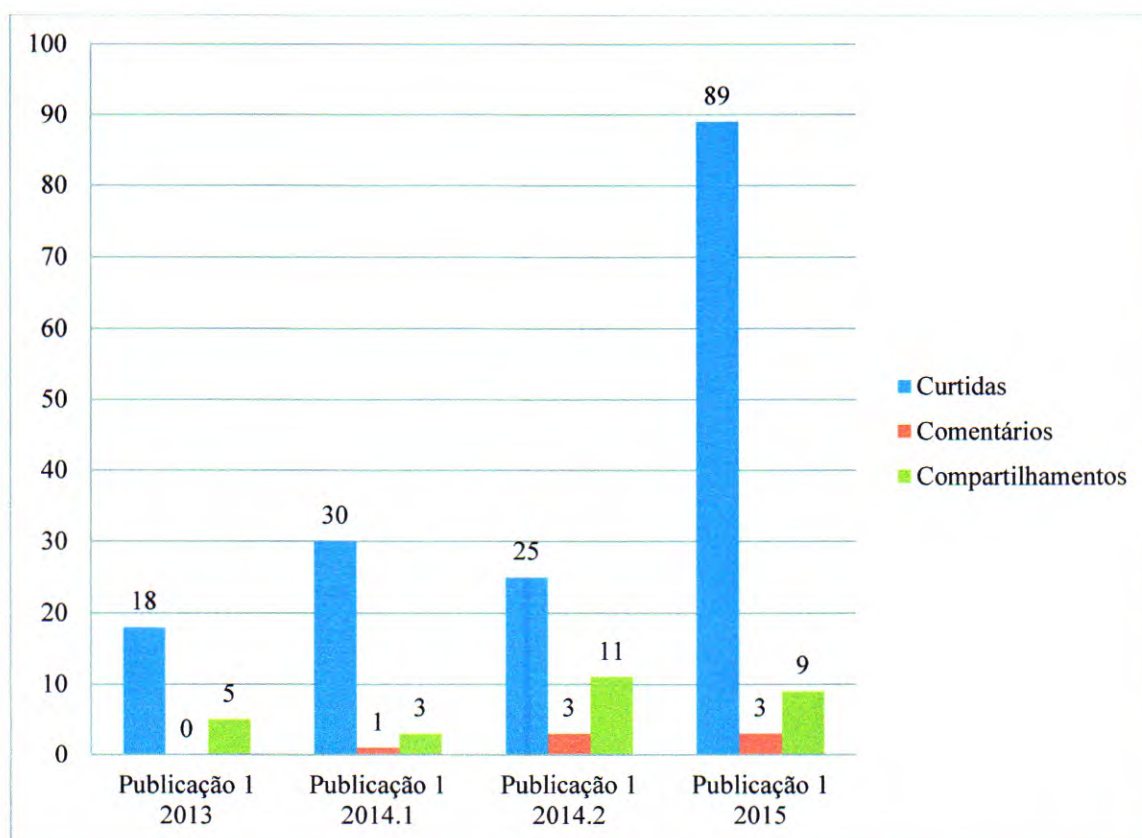
Posteriormente nos Resultados e Discussões é mostrada como é formada a rede, os usuários e publicações com mais autoridade durante o período de existência da página até depois do momento em que as estratégias de marketing são efetuadas e as características da rede explorada.

## 5 RESULTADOS E DISCUSSÃO

Os resultados dos dados explorados e analisados possibilitaram compreender quais são os principais usuários e publicações dessa rede, quais usuários consegue propagar uma informação de forma mais rápida, o alcance das informações e a análise das estratégias de marketing utilizadas.

Foram analisadas as cinco publicações com maior alcance de público nos semestres dos seguintes anos: 2013, 2014 e no mês de janeiro até junho de 2015. Observou-se por meio dos grafos analisados que as publicações com imagens que mostraram as atividades realizadas pela obra social e publicações com mensagens obtiveram um maior alcance.

Ficou perceptível que o horário dessas publicações também influenciou no seu alcance, publicações com horário próximo ao meio dia obtiveram mais interação do público. A Figura 7 mostra as publicações com maior número de envolvimento durante cada espaço de tempo.

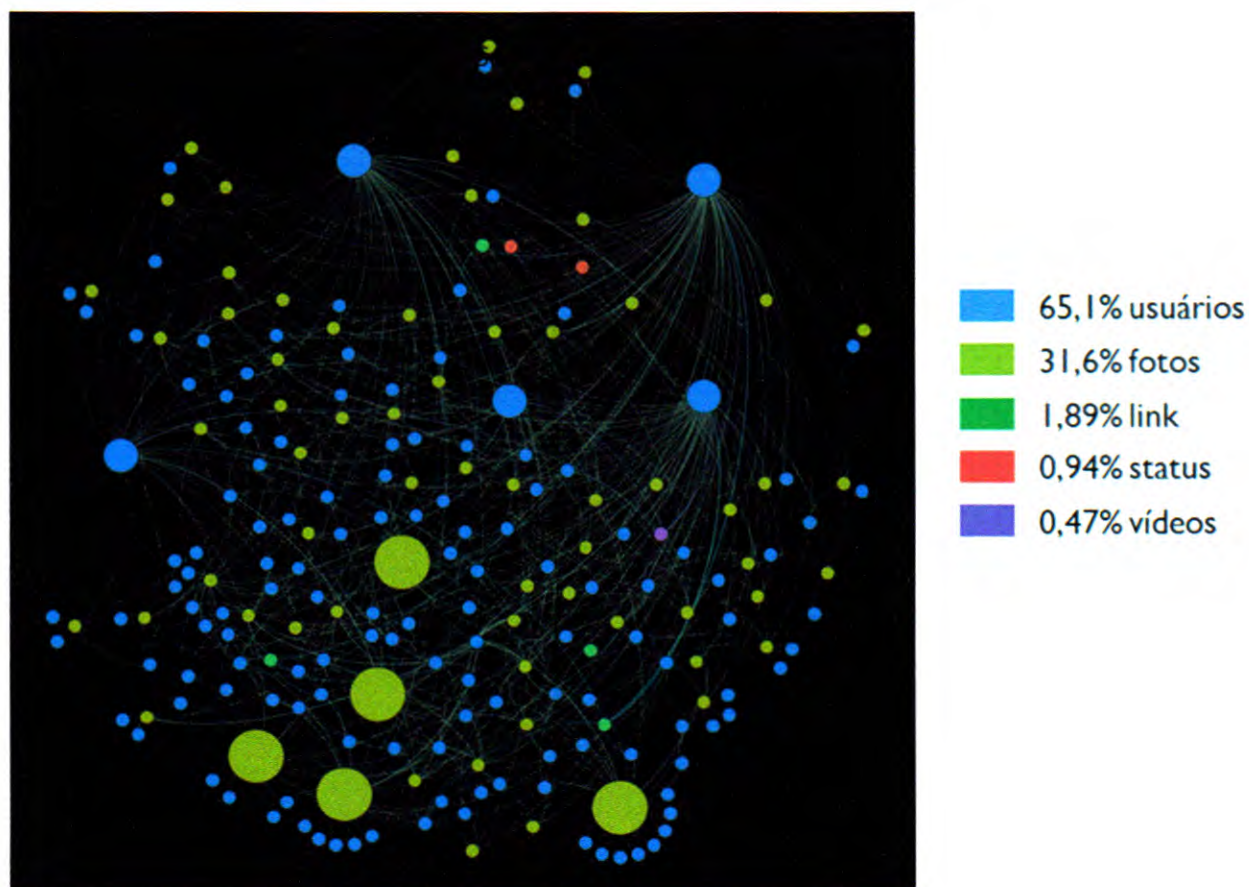


**Figura 3:** Gráfico: Publicações em períodos de tempo diferentes.

Fonte: Próprio Autor (2015).

As estatísticas de grau possibilitaram a visão de todas as conexões da rede, mostrando a interação dos usuários com as publicações da página, o grau médio definiu o peso dos nós em relação à quantidade de suas conexões, por meio desta função foi possível comparar diferentes informações na rede.

Utilizando essas estatísticas de grau e como parâmetro os tipos de nós da rede classificados de acordo com sua divisão e usando o algoritmo *Force Atlas* tornou-se possível equilibrar os nós do grafo e assim foi gerada a Figura 8 que é um grafo sobre nós da rede no primeiro semestre de 2014, contendo 212 nós sendo usuários e publicações, na imagem os cinco nós maiores são as publicações com fotos que obtiveram o maior alcance, outros cinco nós azuis em destaque pelo tamanho são os usuários que mais interagiram nesse período e os outros sem destaques pelo tamanho são publicações e usuários.



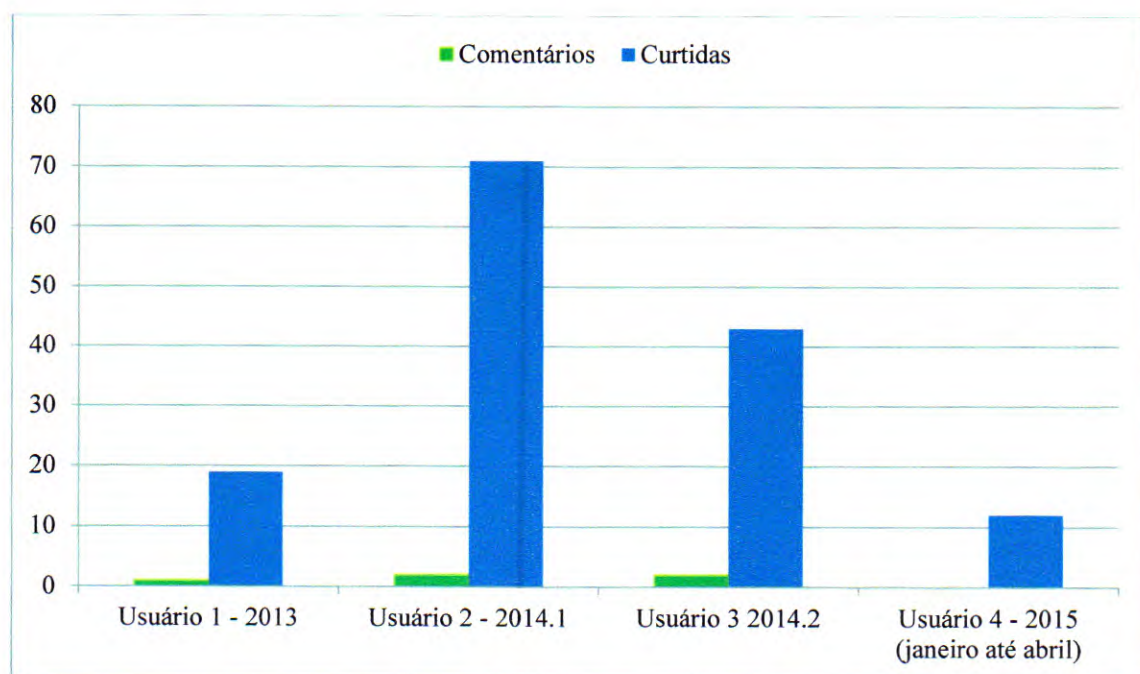
**Figura 4:** Grafo de divisão dos nós da rede no período de 2014.1 (*Gephi*).

Fonte: Próprio Autor (2015).



Foi identificado o usuário que mais se envolveu com as publicações em cada período de tempo, a partir disso é observado na Figura 9, o período de tempo que obteve mais interação do público. Foi verificado que o usuário com mais influencia nos três primeiros períodos é o mesmo, em relação ao período referente ao ano de 2015 percebeu-se que um usuário da rede social conseguiu ter maior influência.

Os usuários mais atuantes na rede conseguem influenciar outros membros do *Facebook* por meio da propagação de informação, portanto estas pessoas são essenciais na divulgação direta ou indireta da página, contribuindo com o possível alcance de mais colaboradores dispostos a aderir à instituição.

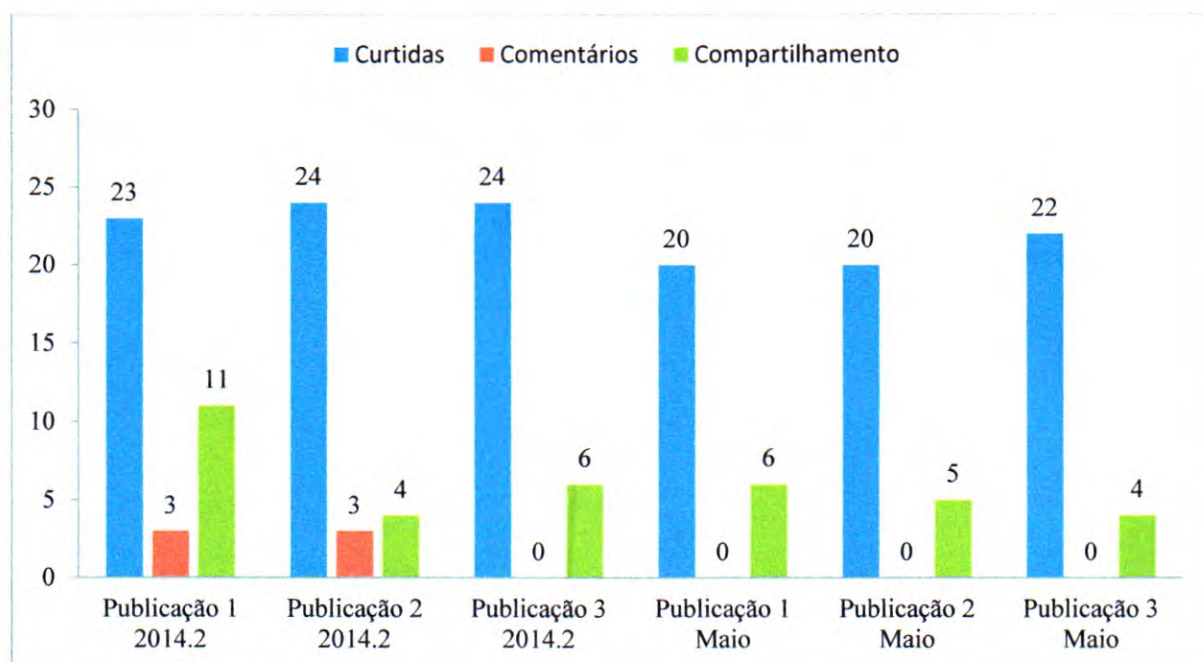


**Figura 5:** Gráfico: Envolvimento dos usuários.

Fonte: Próprio Autor (2015).

Nos primeiros quatro meses do ano de 2015 a página já conseguia um alcance maior nas publicações, com base nos dados analisados e buscando aumentar o público da página foram propostos os seguintes métodos: o aumento de postagens semanais em horários em que foi verificado o aumento de envolvimento dos usuários e a realização de uma promoção na página, viabilizando o aumento da interação do público com a página e o crescimento no número de colaboradores na instituição. Ao comparar os períodos antes da realização de propostas de marketing tornou-se notável o aumento de seguidores na página e consequentemente o aumento de usuários que interagem com as publicações da página.

A Figura 10 mostra as três publicações com maior alcance do segundo semestre de 2014 e no mês de maio de 2015, no qual foi aplicada a estratégia de marketing para publicações semanais em horários específicos. É visível que a interação de usuários em maio aumentou quase alcançando os mesmos números de envolvimento do público do segundo semestre de 2014.



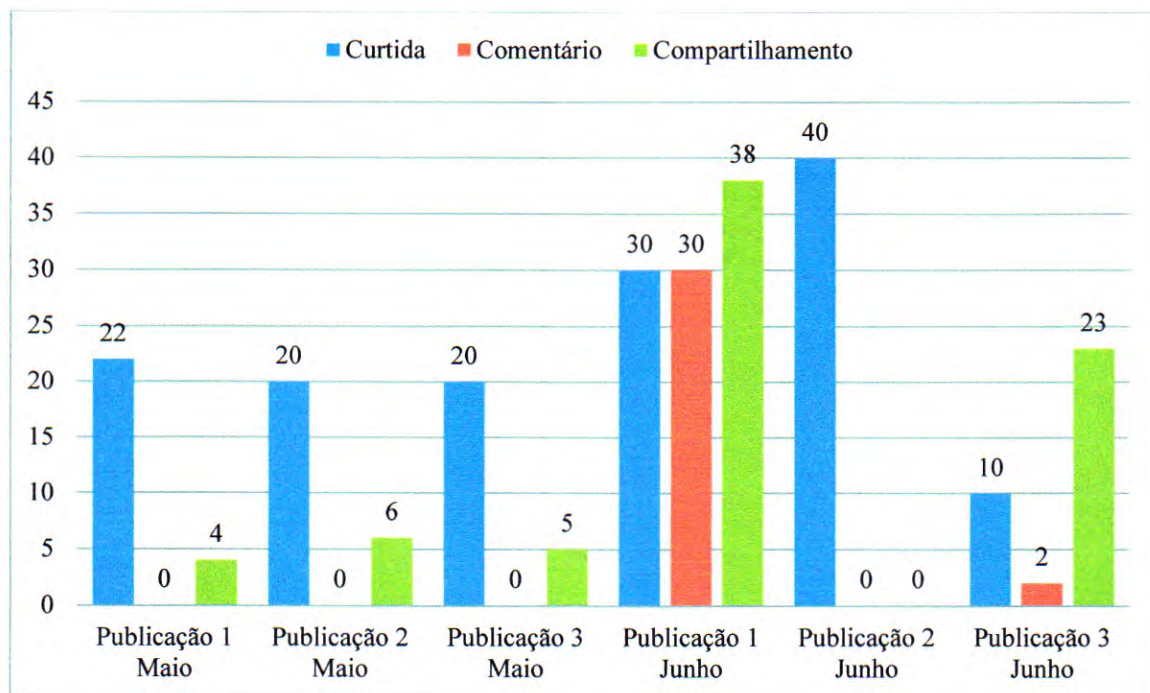
**Figura 6:** Gráfico: Publicações do segundo semestre de 2014 e maio de 2015.

Fonte: Próprio Autor (2015).

Na primeira quinzena do mês de junho de 2015, foi realizada uma promoção na página, no qual os usuários teriam que interagir com determinada publicação para assim participarem do sorteio de um livro, visto que isso iria contribuir no acréscimo de curtidas da página, desta forma aumentando o público.

Na Figura 11 é possível ver a comparação das três publicações com maior alcance nos meses de maio e junho, assim sendo notável que no mês de junho as publicações obtiveram um maior alcance.





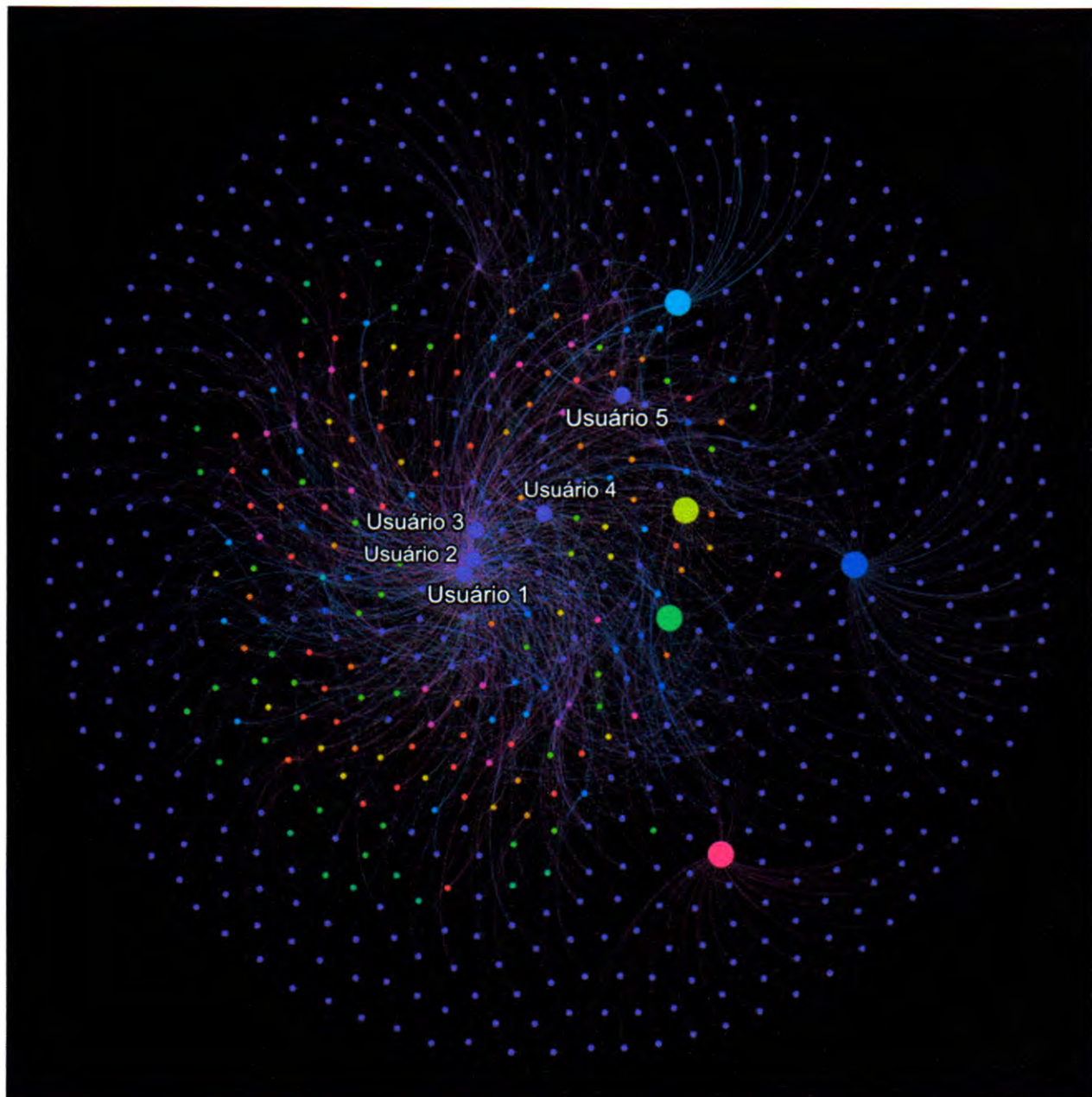
**Figura 7:** Gráfico: Comparação das publicações nos meses de maio e junho.

Fonte: Próprio Autor (2015).

Os dados desde o início da página até a primeira quinzena de junho de 2015 foram analisados em conjunto, desta forma conseguimos identificar as cinco publicações com maior alcance e os cinco usuários que mais propagaram informações.

Na Figura 12 temos um grafo utilizando como grau de entrada o algoritmo HITS (Hyperlink-Induced Topic Search) e o algoritmo de layout Fruchterman Reingold que distribuiu os nós de forma igualitária, diminuindo o cruzamento de arestas e proporcionando simetria. O grafo também foi personalizado com outras funções, como por exemplo, destacando os nós que possuem maior importância.

Os usuários e publicações que tiveram o maior número de conexões ficaram próximos ao centro, portanto isso permitiu descobrir os usuários e publicações que tiveram o maior número de interações, desta forma ficaram em destaque pelo tamanho do nó as cinco publicações que conseguiram mais interação do usuário, ou seja, em que o grau de entrada foi maior, sendo destacados também os cinco usuários com maior grau de saída, assim possuindo o envolvimento mais forte na página.

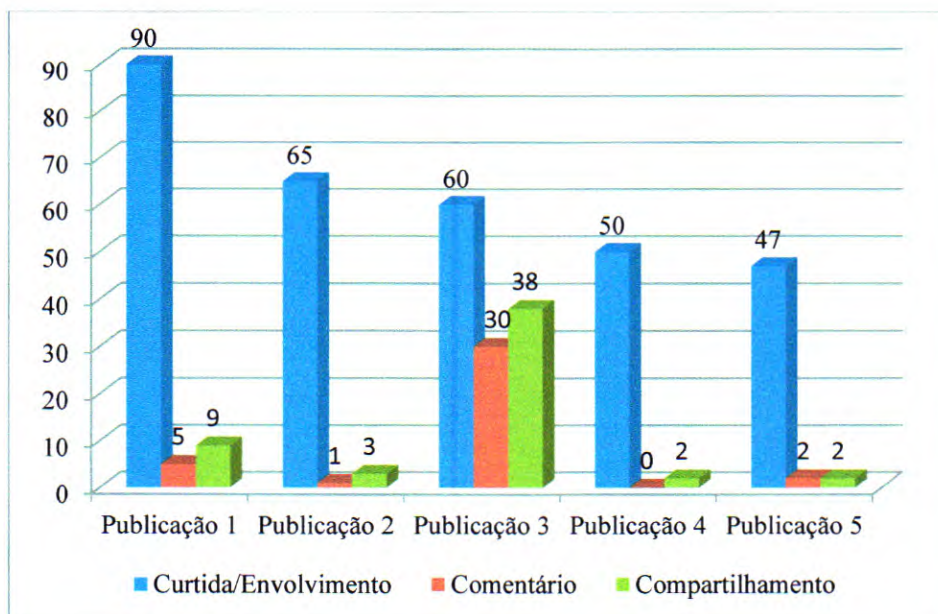


**Figura 8:** Grau de entrada (*Gephi*).

Fonte: Próprio Autor (2015).

Na análise dos dados conforme a Figura 13 foi visto que a publicação com maior alcance é referente a um álbum de fotos sobre atividades realizadas na instituição publicada no dia 07 de março de 2015, a segunda publicação com maior alcance é referente à promoção lançada em 02 de junho de 2015, as outras três são relacionadas à divulgação, atividades e convite para eventos realizados pela instituição no mês de fevereiro e abril de 2015.

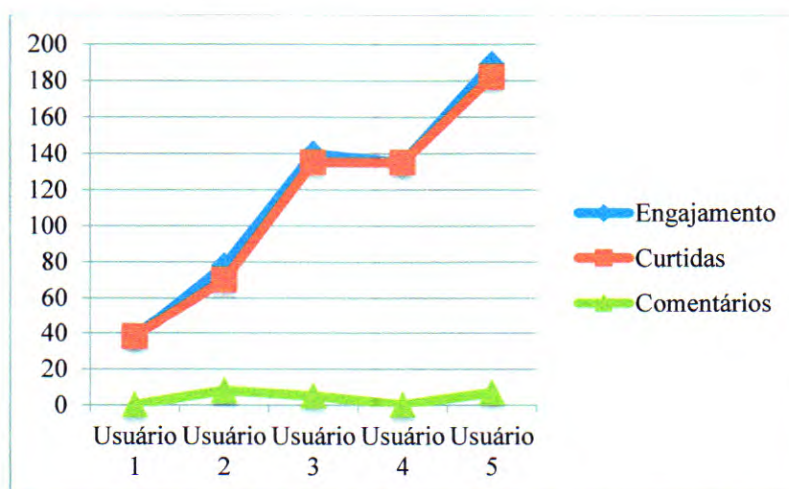




**Figura 9:** Gráfico: Publicações com maior alcance.

Fonte: Próprio Autor (2015).

A Figura 14 mostra os cinco usuários com mais influencia ao longo de quase dois anos da página Obras Sociais Luz da Esperança, é perceptível que a maior interação dos usuários é realizada em forma de curtidas em postagens, na análise dos dados foi possível descobrir que alguns destes usuários atuaram como influenciadores em diferentes períodos. Foi observado que o usuário com maior influencia é do sexo masculino e os outros usuários são do sexo feminino, estas cinco pessoas possuíram uma função importante, pois devido à interatividade com a página conseguiram propagar informações sobre a instituição e assim alcançar novos curtidores para a página.

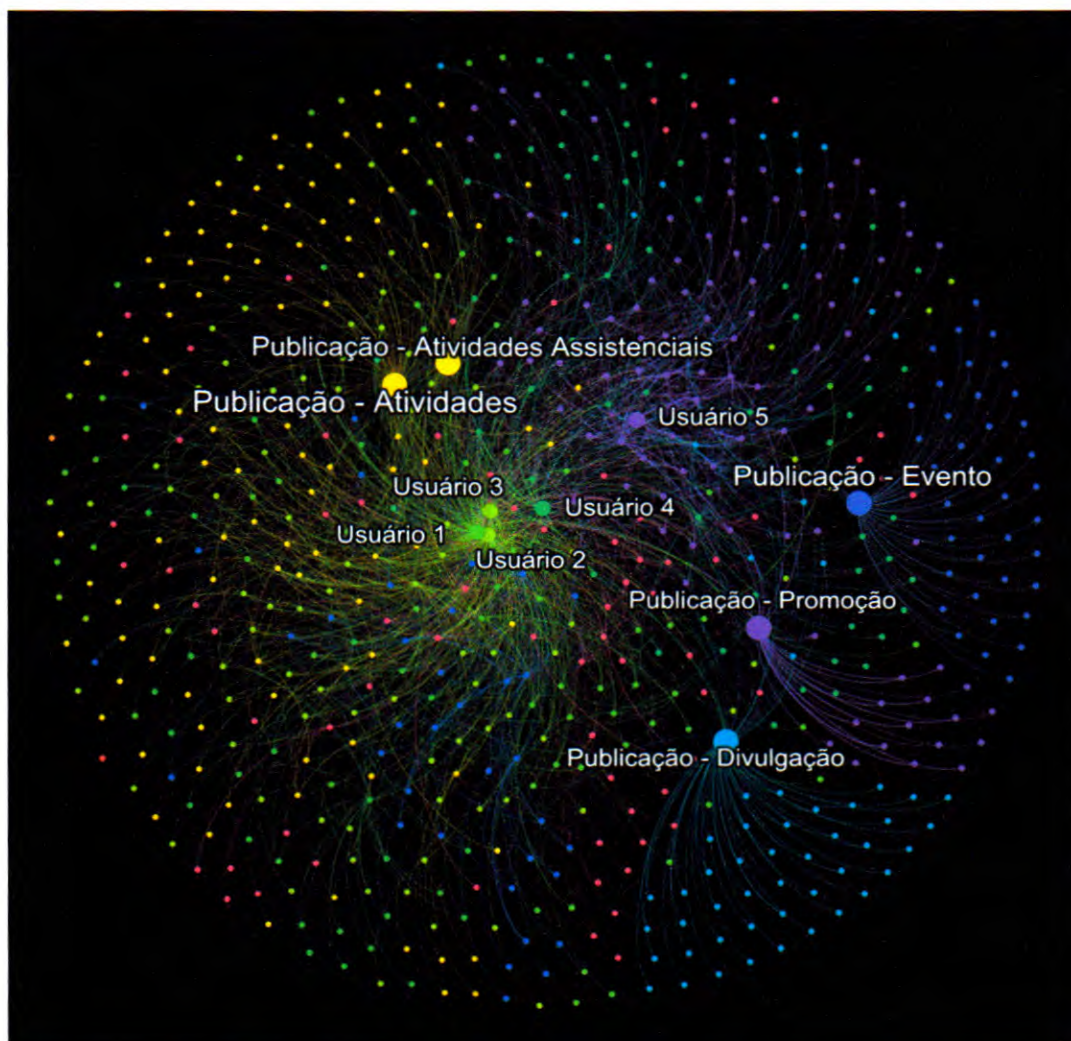


**Figura 10:** Gráfico: Usuários mais influentes.

Fonte: Próprio Autor (2015).

De acordo com Andrade (2013) a centralidade de autovetor tornou possível encontrar os nós que estão mais próximos dos demais, nessa função a importância do nó é baseada nas conexões, é uma medida importante para analisar e propagar informação, quando um nó se conecta a um grande número de nós, se torna uma importante fonte de informação na rede, ou seja, um usuário que interagiu com muitas publicações contribuiu na divulgação da instituição.

A modularidade mostra como a rede é constituída por comunidades, assim formando um conjunto de nós que estão mais conectados entre eles do que a outro nó da rede, consequentemente agrupando em comunidade as publicações da página, desta forma demonstra as publicações que conseguiram um poder de comunicação maior. (ANDRADE, 2013).



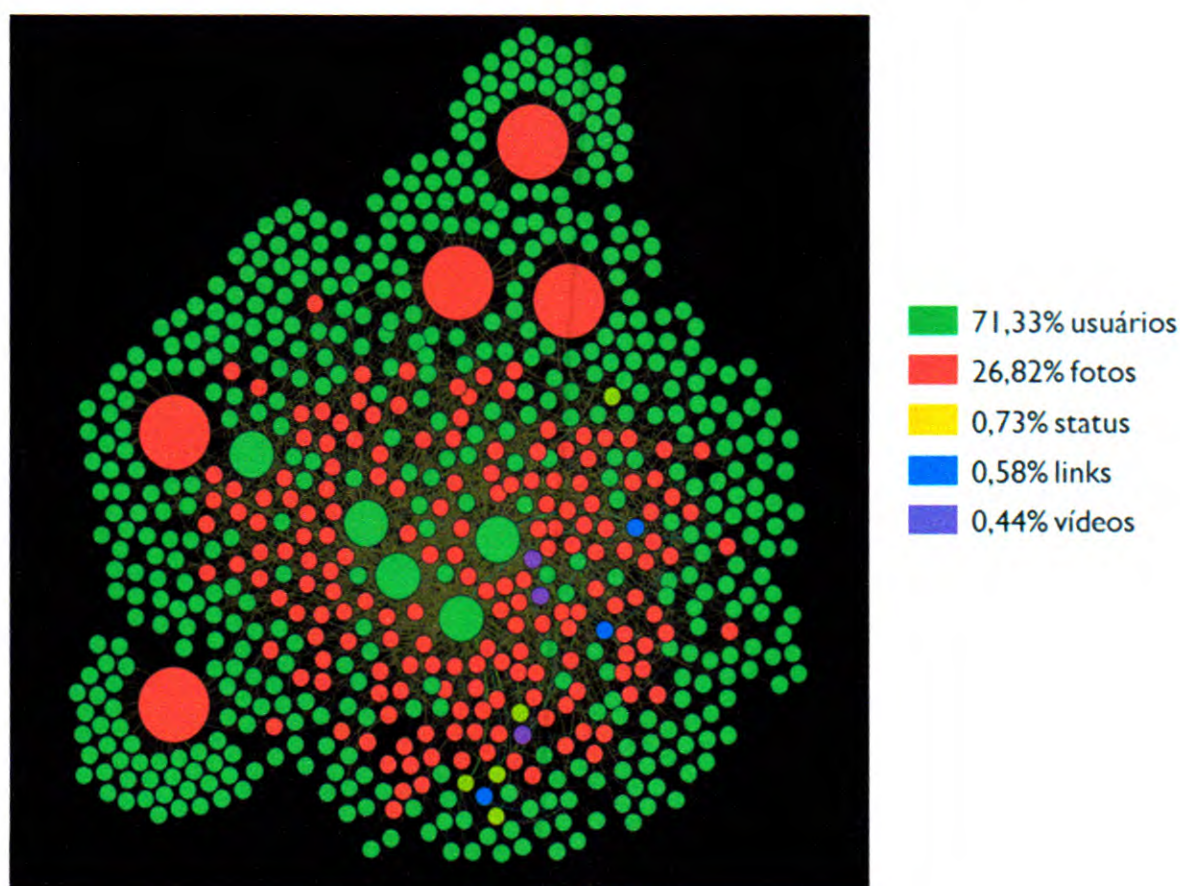
**Figura 11:** Modularidade (*Gephi*).

Fonte: Próprio Autor (2015).



Na Figura 15 utilizando as funções de centralidade de autovetor, modularidade, e o algoritmo Fruchterman Reingold, tornou-se possível observar que os nós foram agrupados de acordo com publicações em determinado período de tempo, sendo que a cor do nó usuário foi determinada conforme a cor da primeira publicação em que ele teve interação na página, na análise do grafo foi possível perceber nós sem nenhuma entrada, no qual são publicações em que não houve nenhum envolvimento de usuários.

Por meio do algoritmo *Force Atlas* foi possível comparar a posição de um nó com os outros, a figura 16 mostra os tipos de nós da rede, sendo que 71,33% é formada por usuários (verde), 26,82% fotos (vermelho), 0,73% status (amarelo), 0,58% links (azul) e 0,44% vídeos (lilás). É notável que os usuários que mais interagiram ficaram no centro do grafo, enquanto os que se envolveram pouco ficaram mais afastados do centro, as publicações que possuíram alcance maior ficaram mais cercadas de um grupo de usuários, foram destacadas no grafo as cinco publicações e cinco usuários mais importantes na rede.



**Figura 12:** Tipos de nós (*Gephi*).

Fonte: Próprio Autor (2015).

## 6 CONSIDERAÇÕES FINAIS

A análise dos dados se mostrou eficaz, por meio de comparações de períodos diferentes foi possível perceber o tipo de publicação que obteve maior alcance, o horário melhor para publicar e obter um alcance maior de público, os usuários que possuem um maior poder de influencia que se tornaram fundamentais para propagar informações. A utilização do software *Gephi* foi de grande importância, pois proporcionou explorar o conhecimento por intermédio do laboratório de dados e em formato de grafos.

Um ponto a ser observado, foi que devido a API do *Facebook* limitar a exploração de dados, não foi possível ter acesso sobre a conexão entre os usuários por meio dos dados extraídos com o *Netvizz*, sendo que todos os seguidores da página não estão presentes no grafo, pois o aplicativo permitiu apenas coletar os usuários que se envolveram em algum momento com a página, entretanto foi visto que um usuário que interage muito com alguma publicação seja por curtidas, comentários ou compartilhamento, propaga informações sobre a página com seus amigos da rede social, pois suas interações são mostradas no *feed* de notícias de seus amigos. Conclui se que a análise em redes sociais torna viável a utilização de estratégias de marketing com base nos dados explorados para assim ampliar o público de determinada página, no caso de instituições podendo assim aumentar o número de colaboradores.

Pode-se realizar estratégias individualizadas de marketing, já que conseguimos descobrir qual usuário possui maior interação com a página e, desse modo, incentivar ainda mais na divulgação da empresa (instituição) visando um público ainda maior.

Este trabalho pode ser prosseguido por meios de algumas sugestões para futuras melhorias, algumas alternativas em relação às estratégias de marketing e ao uso da tecnologia aplicada.

Primeiramente, a exploração de conhecimento poderia contribuir na descoberta de usuários principais e publicações com maior desempenho na divulgação de um financiamento coletivo da instituição com a finalidade de alcançar mais colaboradores. Uma segunda alternativa seria desenvolver um software que permita uma exploração mais ampla dos dados, possibilitando observar a interação de usuários com outros ligados a mesma rede, tornando viável à pesquisa por *hashtags*(#) na rede social e mostrando informações sobre curtidas desde o tempo de existência da página.



Por fim, outra proposta seria em relação ao marketing, após a identificação de usuários que interagem desde o início da página, estes poderiam receber algum brinde em forma de agradecimento, desta forma iria fortalecer o laço entre a instituição e o usuário, além de mostrar a importância que a pessoa em questão teve no crescimento da página na rede social.

## REFERÊNCIAS

ALMEIDA, L. M.; PADILHA, P. P.; OLIVEIRA, F.L.; PREVIERO, C.A. **Uma Ferramenta para Extração de Padrões**. PALMAS, 2013.

ANDRADE, Josmar; MAZZON, José A.; KATZ, Sérgio. **Boca-a-Boca Eletrônico: explorando e integrando conceitos de marketing viral, buzz marketing e word-of-mouse**. In: ENCONTRO DE MARKETING DA ANPAD. Rio de Janeiro, 2006.

ANDRADE, Priscilla Calmon. **Narrativas controversas: As tramas emergentes da ciberguerra do #Wikileaks**. Dissertação (TCC - Comunicação Social) – Universidade Federal do Espírito Santo, Vitória, 2013.

AZEVEDO, Tatiana Barbosa. **Análise de redes sociais através de softwares de ARS e de data mining: Um estudo de caso em turmas de graduação**. Dissertação (Mestrado de Sistema de Gestão) - Universidade Federal Fluminense, Niterói, 2011.

BACKER, Eric. **Computer Assisted Reasoning in Cluster Analysis**. Prentice Hall, New York, 1995.

BARICHELO, E.M.M.R.; OLIVEIRA, C.C. **O marketing viral como estratégia publicitária nas novas ambiências midiáticas**. 2006

BARTOLOMEU, T. A. **Modelo de investigação de acidentes do trabalho baseado na aplicação de tecnologias de extração de conhecimento**. Dissertação (Pós-Graduação em Engenharia de Produção) – Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina, 2002.

BASTIAN, M.; HEYMAN, S.; JACOMY, M. **Gephi: an open source software for exploring and manipulating networks**. International AAAI Conference on Weblogs and Social Media, 2009.

TOMAËL, M. I., ALCARÁ, A. R., CHIARA, I. G. **Das redes sociais à inovação**. Brasília, 2005.

BENEVUTO, F., ALMEIDA, J.M., SILVA, A.S. **Explorando redes sociais online: Da coleta e análise de grandes bases de dados às aplicações**. In XXIX Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos, Mato Grosso do Sul, 2011.

BRUZZESE, D., BUONO, P. **Combining Visual Techniques for Association Rules Exploration**. In: Proceedings of the working conference on Advanced Visual Interfaces – ACM, Itália, 2004.

CASTELLS, Manuel. **A Galáxia Internet: reflexões sobre internet, negócios e sociedade**. Lisboa: Fundação Calouste Gulbenkian, 2004.

CINTRA, F.C. **Marketing digital: A era da tecnologia on-line**. Investigação, v. 10, n. 1, p. 6-12, 2010.

**Facebook para empresas.** Disponível em:

<<https://www.facebook.com/business/products/pages>> Acesso em: 30 de Março de 2015.

FAYYAD, U. M.; Uthurusamy, R. **Evolving Data Mining into Solutions for Insights.** Communications of the ACM: vol.45, N° 8, 2002.

FAZITO, D. **A análise de redes sociais (ARS) e a migração: Mito e realidade.** In: XIII Encontro da Associação Brasileira de Estudos Populacionais. Minas Gerais, 2002.

MARTINS, A., MORALES, A. T., PACHECO, R. C. S. **Data mining: Transformando dados em informação.** Florianópolis, SC, 1999.

FERREIRA, Gonçalo Costa. **Redes Sociais de Informação: uma história e um estudo de caso.** Perspectivas em da informação. 2011, vol.16, n.3, pp. 208-231.

FREITAS, A. A. **Uma Introdução a Data Mining.** Informática Brasileira em Análise, Centro de Estudos e Sistemas Avançados do Recife (C.E.S.A.R.), Recife, 2000.

FREITAS, C. M. D. S., NEDEL L.P., GALANTE, R.; LAMB L. C.; SPRITZER, A.S.; FUJII, S.; OLIVEIRA, J.P.M.; ARAUJO, R.M.; MORO, M.M. **Extração de conhecimento e análise visual de redes sociais.** In SEMISH – Seminário Integrado de Software e Hardware, Belém do Pará, 2008.

GEPHI. Disponível em: <<https://gephi.github.io/features>> Acesso em: 27 de março de 2015.

GOEBEL, M., GRUENWALD, L. **A survey of data mining and knowledge discovery software tools.** 1999.

HINNEBURG, A., KEIM, D., WAWRYNIUK, M. **HD-Eye: Visual Mining of High-Dimensional Data,** 1999.

KREMER, Ricardo. **Sistema de apoio à decisão para previsões genéricas utilizando técnicas de Data Mining.** Trabalho de conclusão de curso (Graduação) – Curso de Ciências da Computação, Bacharelado, Universidade Regional de Blumenau, Santa Catarina, 1999, p. 18.

KOTLER, Philip. **Administração de marketing.** 10.ed. São Paulo: Prentice Hall, 2000.

LAS CASAS, A. L. **Administração de marketing: conceitos, planejamento e aplicações à realidade brasileira.** São Paulo: Atlas, 2006.

LEITE, N. **Proposta de uma aplicação web para monitorização do impacto de notícias nas redes sociais Facebook e Twitter.** Revista Comunicando, vol. 2, 2013.

METZ, J.; CALVO, R.; SENO, E. R. M.; ROMERO, R. A. F.; LIANG, Z. **Redes Complexas: Conceitos e Aplicações.** Instituto de Ciências Matemáticas e de Computação. São Carlos, 2007.

MIRANDA, H. S.; BRITO, P. F.; FAGUNDES, F. **União Rede Social e Marketing Viral para ampliação da concorrência em processo seletivo (vestibular) de uma Instituição de Ensino Superior.** In: XI Encontro de Estudantes de Informática do Tocantins, 2009.

**NETVIZZ.** Disponível em: <<https://apps.facebook.com/netvizz>> Acesso em: 27 de março de 2015.

OKADA, S.I.; SOUZA, E.M.S. **Estratégias de marketing digital na era da busca.** Revista Brasileira de Marketing, 2011, p. 47.

PANTELI, Niki. **Virtual Social Networks: Mediated, Massive and Multiplayer Sites.** Palgrave-Macmillan, Hampshire, UK, 2009.

PETTENATI, Maria Chiara; RANIERI, Maria; (2006). **Informal learning theories and tools to support knowledge management in distributed CoPs.** In Innovative Approaches for Learning and Knowledge Sharing, 2006.

POLONI, K. M.; TOMAÉL, M.I. **Coleta de dados em plataformas de redes sociais: Um estudo de aplicativos.** WORKSHOP DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 2014, p. 5 – 7.

RECUERO, Raquel. **Redes Sociais na Internet.** Porto Alegre: Sulina, 2009  
DINIZ, C. A. R.; LOUZADA NETO, F. **Data mining: uma introdução.** São Paulo: ABE, 2000.

REZENDE, S. O. **Mineração de Dados.** In: Encontro nacional de inteligência artificial e computacional (ENIAC), São Leopoldo: UNISINOS, 2005.

RIEDER, B. **Studying Facebook via Data Extraction: The Netvizz Application.** University of Amsterdam, 2013.

ROSA, A. C., SILVA, B. D., SILVA, P. L. **Análise de redes sociais aplicadas à engenharia social.** Faculdade de Tecnologia de Guaratinguetá, 2002.

ROSEN, Emanuel. **Marketing boca a boca: como fazer com que os clientes falem de sua empresa, seus produtos e serviços.** São Paulo: Futura, 2001.

SANTOS, M. **Os haters políticos: Monitoramento e mapeamento dos hubs de oposição e do discurso de ódio no Facebook.** Rio de Janeiro, 2014.

SFERRA, Heloisa Helena; CORREA, Ângela M. C. Jorge. **Conceitos e Aplicações de Data Mining.** Revista Ciência & Tecnologia, 2003.

SOUKUP, T., DAVIDSON, I., **Visual Datamining - Techniques and Tools for Data Visualization and Mining,** Ed. Wiley Publishing, 2002.

TAN, P.N., STEINBACH, M., KUMAR, V. **Introduction to data mining.** Boston, 2006.

VAKALI, A. **Evolving social data mining and affective analysis methodologies, framework and applications.** 2012.

WANG, Chia-Chen et al. **Why Focal Firms Share Information? A Study of the Effects of Power and Information Technology Competence.** In: PACIFIC ASIA CONFERENCE ON INFORMATION SYSTEMS (PACIS). 2008.

WASSERMAN, S.; FAUTS K. **Social Network Analys Methods and Applications**. Cambridge University Press, 1994.

YAMAGUCHI, J.K. **Diretrizes para a escolha de técnicas de visualização aplicadas no processo de extração do conhecimento**. Dissertação (Pós-Graduação em Ciência da Computação) – Universidade Estadual de Maringá, Paraná, 2010.