

Marina Carradore Sérgio

**UM MODELO BASEADO EM ONTOLOGIA E ANÁLISE DE
AGRUPAMENTO PARA SUPORTE À GESTÃO DE IDEIAS**

Dissertação submetida ao Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento da Universidade Federal de Santa Catarina para a obtenção do Grau de Mestre em Engenharia e Gestão do Conhecimento.

Orientador: Prof. Dr. Alexandre Leopoldo Gonçalves.

Coorientador: Prof. Dr. João Artur de Souza.

Florianópolis
2016

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Sérgio, Marina Carradore

Um Modelo Baseado em Ontologia e Análise de Agrupamento
para Suporte à Gestão de Ideias / Marina Carradore Sérgio ;
orientador, Alexandre Leopoldo Gonçalves ; coorientador,
João Artur de Souza. - Florianópolis, SC, 2016.
128 p.

Dissertação (mestrado) - Universidade Federal de Santa
Catarina, Centro Tecnológico. Programa de Pós-Graduação em
Engenharia e Gestão do Conhecimento.

Inclui referências

1. Engenharia e Gestão do Conhecimento. 2. Gestão de
Ideias. 3. Análise de Agrupamento. 4. Ontologia. 5.
Inovação. I. Gonçalves, Alexandre Leopoldo . II. Souza, João
Artur de. III. Universidade Federal de Santa Catarina.
Programa de Pós-Graduação em Engenharia e Gestão do
Conhecimento. IV. Título.

Marina Carradore Sérgio

UM MODELO BASEADO EM ONTOLOGIA E ANÁLISE DE AGRUPAMENTO PARA SUPORTE À GESTÃO DE IDEIAS

Esta Dissertação foi julgada adequada para obtenção do Título de “Mestre em Engenharia e Gestão do Conhecimento”, e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento.

Florianópolis, 23 de Fevereiro de 2016.

Prof. Roberto Carlos dos S. Pacheco, Dr.
Coordenador do Curso

Banca Examinadora:

Prof. Alexandre Leopoldo Gonçalves, Dr.
Orientador
Universidade Federal de Santa Catarina

Prof. João Artur de Souza, Dr.
Coorientador
Universidade Federal de Santa Catarina

Prof.^a Gertrudes Aparecida Dandolini, Dr.^a
Universidade Federal de Santa Catarina

Prof. Denilson Sell, Dr.
Universidade Federal de Santa Catarina

Prof. Roderval Marcelino, Dr.
Universidade Federal de Santa Catarina

Este trabalho é dedicado a todos que
direta ou indiretamente contribuíram
em minha formação acadêmica.

AGRADECIMENTOS

Ao longo desta jornada muitos foram os desafios encontrados, porém com o apoio e a compreensão destas pessoas foi possível alcançar o término desta jornada.

Agradeço ao meu orientador, Dr. Alexandre Leopoldo Gonçalves, pelo apoio e dedicação. Pelos momentos de compreensão, conversas, discussões, que sempre muito agregaram à minha formação. Sua participação foi fundamental para que os objetivos fossem alcançados.

Aos professores Dra. Gertrudes Aparecida Dandolini, Dr. Denilson Sell e Dr. Roderval Marcelino, por aceitarem o convite para a banca de defesa. Uma honra tê-los como avaliadores deste trabalho.

Ao Gustavo, por estar presente em todos os momentos me apoiando e me compreendendo, com certeza esta jornada seria difícil sem você.

À minha família, pelas palavras de afeto e pelos momentos com os quais não pude compartilhar.

A Deus, a quem devo o dom da vida.

Ao Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, pela oportunidade de participar desse prestigiado curso.

Agradeço, também, a todos que contribuíram com o desenvolvimento desta dissertação: aos colegas, amigos e professores. Obrigada pelo incentivo e colaboração no andamento e finalização deste trabalho.

“A mente que se abre a uma nova ideia jamais
voltará ao seu tamanho original.”

(Albert Einstein)

RESUMO

Na sociedade do conhecimento o processo de inovação é visto como um desafio às organizações que o utilizam como um diferencial para obtenção de vantagem competitiva frente ao mercado. Ao gerenciar e alinhar o processo de inovação às estratégias organizacionais torna-se possível proporcionar aos clientes produtos novos ou melhores, ou ainda inovar em processos ou em *marketing*. A Gestão de Ideias integra este processo e apresenta-se como um fator primordial para o sucesso do mesmo. Pode ser vista como uma área potencial que se encontra em expansão, possibilitando gerar retorno financeiro para a organização. O principal objetivo da área está em organizar ideias e implementá-las de maneira eficaz e eficiente. Porém, devido ao número de ideias coletadas em uma organização, o processo de seleção de ideias para execução torna-se não trivial. Ao utilizar técnicas de agrupamento é possível reunir grupos de ideias semelhantes e facilitar a visualização das informações. Esta dissertação possui como objetivo apresentar um modelo baseado em ontologia e na análise de agrupamento para suporte à Gestão de Ideias, visando auxiliar no processo de tomada de decisão. Para demonstração de viabilidade do modelo proposto, foi desenvolvido um protótipo para suportar as fases de indexação, extração, validação, agrupamento e visualização das ideias. O protótipo foi aplicado em três cenários de estudo utilizando ideias coletadas nos sites das empresas Starbucks® e Dell®. A partir da aplicação do protótipo identificou-se que ao analisar grupos de ideias semelhantes, e não somente ideias isoladamente, estes podem indicar padrões e tendências que auxiliem na tomada de decisão de quais ideias deveriam ou não serem implementadas. Como resultado apresenta-se um modelo capaz de auxiliar na tomada de decisão e no processo de destinação de recursos para investimentos em ideias. As evidências obtidas a partir da análise dos textos providos pelos especialistas da organização comprovam as descobertas efetuadas pelo modelo.

Palavras-chave: Gestão de Ideias. Análise de Agrupamento. Ontologia. Inovação.

ABSTRACT

In the knowledge society the innovation process is challenging to organizations that use it as a differential to obtain competitive advantage over the market. Managing and aligning the innovation process to organizational strategies make possible to provide customers with better or new products, innovating in processes or marketing. The Ideas Management integrates this process and presents itself as a major factor in its success. It can be seen as a potential area, which is expanding, allowing generate financial return for organizations. The main objective of the area is to organize ideas and implement them effectively and efficiently. However, due to the number of ideas collected in an organization the process of selection of ideas for execution becomes non-trivial. Using clustering techniques can compose groups of similar ideas as well as to make easy information visualization. This work has intends to present a model based on ontology and cluster analysis to support the Idea Management, aiming to help in the decision-making process. To demonstrate the feasibility of the proposed model, a prototype was developed to deal with the phases of indexing, extraction, validation, clustering and visualization of ideas. The prototype was applied in three study scenarios using ideas collected on the websites of Starbucks® and Dell® companies. From the application of the prototype it was found that when analyzing of like-minded groups, not just ideas alone, these can indicate patterns and trends to assist in making decisions on which ideas should or should not be implemented. As a result, the model is able to assist in decision making and in the process of allocation of funds for idea investments. Evidence obtained from the analysis of the texts provided by the organization's experts confirm the findings made by the model.

Keywords: Idea Management. Cluster Analysis. Ontology. Innovation.

LISTA DE FIGURAS

Figura 1 - Procedimentos para realização da dissertação.....	30
Figura 2 - Áreas de concentração envolvidas na pesquisa.....	32
Figura 3 – Benefícios da aplicação do processo de Gestão de Ideias	39
Figura 4 – Modelo <i>Stage Gate</i>	42
Figura 5 – Metodologia de criação de ideias	42
Figura 6 – Modelo de planejamento da inovação	43
Figura 7 – Túnel de ideias.....	44
Figura 8 – Processo de inovação segundo Crawford e Benedetto	44
Figura 9 – Modelo de inovação segundo Whitney.....	45
Figura 10 – Modelo de <i>front end</i>	46
Figura 11 - Etapas do processo de agrupamento.....	56
Figura 12 - Abordagem aglomerativa e abordagem divisiva	58
Figura 13 - Modelo proposto	68
Figura 14 – Formato de exposição das ideias no site da empresa Starbucks® ..	69
Figura 15 - Documento XML	70
Figura 16 – Estrutura do objeto de JSON para indexação	71
Figura 17 - Ontologia de domínio.....	72
Figura 18 – Classes presentes na ontologia de domínio.....	73
Figura 19 – Classe <i>Idea</i> e seus relacionamentos.....	74
Figura 20 – Classe <i>Cluster</i> e seus relacionamentos	74
Figura 21 - Processo de agrupamento	76
Figura 22 – Explicitação do conhecimento por meio de uma rede	77
Figura 23 – Exemplo de uma ideia estruturada no formato XML	80
Figura 24 – Nuvem dos termos mais significativos do primeiro cenário de estudo.....	82
Figura 25 – Agrupamento formado a partir do primeiro cenário de estudo.....	85
Figura 26 – Agrupamento formado a partir da query <i>App</i>	86
Figura 27 – Agrupamento formado a partir da query <i>Windows phone</i>	87
Figura 28 – Nuvem dos termos mais significativos do segundo cenário de estudo.....	89
Figura 29 – Agrupamento formado a partir do segundo cenário de estudo	89
Figura 30 – Visualização do nodo <i>Starbucks app</i>	90
Figura 31 – Visualização do nodo <i>Windows phone</i>	91
Figura 32 – Visualização do nodo <i>Windows app</i>	92
Figura 33 – Visualização do nodo <i>Android app</i>	92
Figura 34 – Visualização do nodo <i>Drink orders</i>	93
Figura 35 – Visualização do nodo <i>Coffee cups</i>	94
Figura 36 – Nuvem dos termos mais significativos do terceiro cenário de estudo	96
Figura 37 – Agrupamento formado a partir do terceiro cenário de estudo.....	97
Figura 38 – Visualização do nodo <i>Computer screen</i>	98
Figura 39 – Visualização do nodo <i>Desktop computer</i>	98
Figura 40 – Visualização do nodo <i>Alienware 17</i>	99

Figura 41 – Visualização do nodo <i>Keys of the keyboard</i>	100
Figura 42 – Anúncio <i>Stabucks Delivery in Seattle</i>	101
Figura 43 – Anúncio <i>Saving energy with LED Lighting</i>	101
Figura 44 – Anúncio <i>Stabucks app + Apple Watch</i>	102
Figura 45 – Anúncio <i>Happy Member Mondays</i>	103
Figura 46 – Anúncio <i>Android App Update</i>	103
Figura 47 – Anúncio <i>Mobile Ordering Coming to Portland, OR and Menu Returning Through the Updated Starbucks® app for iPhone®</i>	104

LISTA DE QUADROS

Quadro 1 - Tipos de ontologias.....	49
-------------------------------------	----

LISTA DE TABELAS

Tabela 1 - Termos mais frequentes no primeiro cenário de estudo.....	83
Tabela 2 - Termos mais frequentes no segundo cenário de estudo	88
Tabela 3 - Termos mais frequentes no terceiro cenário de estudo	95
Tabela 4 – Índice de modularidade obtido com aplicação do algoritmo	96

LISTA DE ABREVIATURAS E SIGLAS

API – *Application Programming Interface*
BD – Banco de Dados (*Database*)
CPU – *Central Processing Unit*
EC – Engenharia do Conhecimento
EGC – Engenharia e Gestão Conhecimento
GC – Gestão do Conhecimento
IA – Inteligência Artificial
IMS – *System Management Idea*
KDD – *Knowledge Discovery in Database*
KDT – *Knowledge Discovery in Text*
LED – *Light Emitting Diode*
LSI – *Latent Semantic Indexing*
OWL - *Ontology Web Language*
PLN – Processamento de Linguagem Natural
RI – Recuperação de Informação
SBC – Sistemas Baseados em Conhecimento
SVD – *Singular Value Decomposition*
SVM – *Support Vector Machines*
T – Documento de Texto
tf-idf – *Term Frequency–Inverse Document Frequency*
TI – Tecnologia da Informação
UFSC – Universidade Federal de Santa Catarina
XML – *eXtensible Markup Language*

SUMÁRIO

1 INTRODUÇÃO	23
1.1 DEFINIÇÃO DO PROBLEMA.....	25
1.2 OBJETIVOS.....	26
1.2.1 <i>Objetivo Geral</i>	27
1.2.2 <i>Objetivos Específicos</i>	27
1.3 JUSTIFICATIVA E RELEVÂNCIA DO TEMA	27
1.4 DELIMITAÇÕES DA PESQUISA	29
1.5 PROCEDIMENTOS METODOLÓGICOS.....	30
1.6 ADERÊNCIA E INTERDISCIPLINARIDADE	31
1.6.1 <i>Identidade</i>	31
1.6.2 <i>Contexto estrutural no EGC</i>	32
1.6.3 <i>Referências Factuais</i>	34
1.7 ESTRUTURA DO TRABALHO.....	35
2 FUNDAMENTAÇÃO TEÓRICA	37
2.1 GESTÃO DE IDEIAS	37
2.1.1 <i>Sistemas de Gestão de Ideias</i>	40
2.1.2 <i>Modelos para Gestão de Ideias</i>	41
2.2 ONTOLOGIA	46
2.2.1 <i>Trabalhos correlatos relacionados ao domínio de Gestão de Ideias e Aplicações Semântica</i>	51
2.3 DESCOBERTA DE CONHECIMENTO	53
2.4 ANÁLISE DE AGRUPAMENTO.....	55
2.4.1 <i>Algoritmos hierárquicos</i>	58
2.4.2 <i>Algoritmos por particionamento</i>	59
2.4.2.1 <i>K-means</i>	60
2.4.2.2 <i>Lingo</i>	60
2.4.3 <i>Índice de modularidade</i>	61
2.4.3.1 <i>Medidas de Centralidade</i>	62
2.4.4 <i>Trabalhos correlatos relacionados ao domínio de Gestão de Ideias e Análise de Agrupamento</i>	64
2.5 CONSIDERAÇÕES FINAIS	66
3 MODELO PROPOSTO	67
3.1 DESCRIÇÃO DO MODELO	67
3.1.1 <i>Construção da estrutura de dados para processamento pelo modelo</i>	68
3.1.2 <i>Indexação das ideias</i>	70
3.1.3 <i>População da ontologia</i>	71
3.1.4 <i>Aplicação do algoritmo de agrupamento</i>	75

3.1.5 Avaliação e explicitação do conhecimento.....	77
3.2 CONSIDERAÇÕES FINAIS.....	78
4 APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS	79
4.1 CENÁRIOS DE ESTUDO.....	79
4.1.1 Preparação dos dados	80
4.1.2 Algoritmo de agrupamento	81
4.1.3 Resultados dos cenários de estudo	82
4.1.3.1 Primeiro cenário de estudo.....	82
4.1.3.2 Segundo cenário de estudo	88
4.1.3.3 Terceiro cenário de estudo	94
4.2 ANÁLISE DE IMPLEMENTAÇÃO DE IDEIAS DA EMPRESA STARBUCKS®	100
4.3 CONSIDERAÇÕES FINAIS.....	105
5 CONSIDERAÇÕES FINAIS.....	107
5.1 CONCLUSÕES	107
5.2 PERSPECTIVAS DE TRABALHOS FUTUROS	108
REFERÊNCIAS	111

1 INTRODUÇÃO

A manutenção da vantagem competitiva por parte das organizações precisa de inovações e do desenvolvimento de produtos que satisfaçam as expectativas dos consumidores. Porém, inovar não é um processo trivial.

Segundo Trott (2012), a inovação compreende o processo de geração de ideias, o desenvolvimento de tecnologias e processos, ou ainda a fabricação e marketing de um produto. Novas ideias podem desencadear o desenvolvimento de novos produtos e o processo de inovação (WITELL; LÖFGREN; GUSTAFSSON, 2011).

O processo de inovação torna-se um quesito desafiador e não trivial para as organizações. Como todo processo, é necessário que ocorra o gerenciamento do mesmo, consistindo fundamentalmente em alinhá-lo as estratégias, aos objetivos e aos valores da organização.

Segundo Barbieri, Álvares e Cajazeira (2009) um dos principais insumos deste processo são as ideias. Ideias compõem a matéria-prima para o processo de inovação (BREM; VOIGT, 2007). Organizações que consideram a inovação como uma estratégia competitiva devem estar preparadas para a criação e gestão de ideias (BARBIERI; ÁLVARES; CAJAZEIRA, 2009).

A Gestão de Ideias pode representar o ponto inicial do processo de inovação e auxiliar na identificação de oportunidades viáveis e plausíveis de serem implementadas pela organização (BJÖRK; BOCCARDELLI; MAGNUSSON, 2010; VANDENBOSCH, SAATCIOGLU; FAY, 2006).

A Gestão de Ideias apresenta-se como um fator crucial para o sucesso do processo de inovação. Segundo Sint *et al.* (2010) e Xie e Zhang (2010), através da Gestão de Ideias é possível aumentar a produtividade das empresas, favorecer a criação de novos produtos e melhorar os processos realizados na empresa, minimizando os custos e mantendo a organização competitiva.

Sistemas de Gestão de Ideias podem auxiliar no processo de administrar, avaliar e selecionar ideias (WESTERSKI; DALAMAGAS; IGLESIAS, 2013). Estes sistemas empregam tecnologias da informação para apoiar o processo de inovação e estão sendo considerados como um ramo promissor do mercado de *software* (FENN; LEHONG, 2011; WESTERSKI; IGLESIAS, 2011).

O desenvolvimento de Sistemas de Gestão de Ideias recebe a contribuição da área de Engenharia do Conhecimento. A Engenharia do Conhecimento contribui na aplicação de teorias, métodos e ferramentas

voltadas à modelagem e representação do conhecimento de domínio (STUDER, BENJAMINS; FENSEL, 1998), fornecendo um conjunto de ferramentas para formalizar e explicitar este conhecimento, de modo a auxiliar nas tarefas intensivas em conhecimento e assim fornecer elementos para a gestão do conhecimento (SCHREIBER *et al.*, 2002).

Para representar e explicitar este conhecimento, a Engenharia do Conhecimento utiliza ferramentas como a ontologia para organizar as estruturas de dados a partir de conceitos e relacionamentos (POLI; OBRST, 2010; SUÁREZ-FIGUEROA *et al.*; 2011).

Segundo Fernandes *et al.* (2011) a existência de ferramentas computacionais para amparar o processo de inovação ainda é pequena, sendo que as tecnologias presentes na Web Semântica, por exemplo as ontologias, representam um grande potencial para área.

A meta-informação existente na ontologia do domínio de Gestão de Ideias contribui no agrupamento de ideias em categorias, na busca e até mesmo avaliação das ideias (HÜSIG; KOHN, 2011). Ao agrupar ideias, especialistas de domínio podem interpretar e visualizar melhor as ideias coletadas na organização (POVEDA; WESTERSKI; IGLESIAS, 2012). Uma meta-análise pode identificar agrupamentos (*clusters*) de ideias semelhantes, apontando para possíveis padrões, tendências e demandas de mercado, ou ainda influenciar na decisão dos especialistas de domínio (MAGNUSSON; NETZ; WÄSTLUND, 2014).

Este trabalho apresenta uma contribuição à área de Engenharia do Conhecimento. Área esta que possui suas origens no final de 1970, onde por meio de métodos, técnicas e ferramentas buscou-se o desenvolvimento de sistemas especialistas, sistemas baseados em conhecimento e sistemas de informação intensivos em conhecimento (SCHREIBER *et al.*, 2002).

Nesta perspectiva, acredita-se que o conhecimento do especialista de domínio pode ser modelado, a fim de prover suporte às tarefas intensivas em conhecimento.

A contribuição desta dissertação se dá por meio do desenvolvimento de um modelo capaz de auxiliar no processo de Gestão de Ideias, de modo que por meio de técnicas de agrupamento, ideias semelhantes possam ser agrupadas. A formação dos grupos de ideias semelhantes pode evidenciar padrões e tendências, assim como, possíveis demandas de mercado.

O modelo foi desenvolvido a partir da aquisição de conhecimento e por meio da sistematização dos conceitos envolvendo a área de Gestão de Ideias, baseado em passos metodológicos que resultaram em uma ontologia para representação do conhecimento de domínio, bem como

em um sistema computacional (nível de protótipo) para realizar o processo de agrupamento.

1.1 DEFINIÇÃO DO PROBLEMA

Na sociedade do conhecimento o aumento da concorrência fez com que as organizações realizassem investimentos em inovações, ofertando aos consumidores novos produtos, visando atingir outros mercados, em busca da manutenção da vantagem competitiva (REID; BRETANI, 2004).

O processo de inovação nas organizações tornou-se fundamental para melhorar a produtividade e aumentar as taxas de crescimento da produção (CHANG; CHEN; WEY, 2007). A capacidade de uma organização em produzir inovações está associada a uma demanda contínua de ideias, com o objetivo de desenvolver novos produtos, serviços e tecnologias (BESSANT *et al.*, 2005).

Para Endesley (2010) o ponto central da inovação relaciona-se com a geração e teste de novas ideias. Vygotsky (2007) salienta que o maior insumo presente na sociedade é o conhecimento e a capacidade de gerar ideias inovadoras.

A Gestão de Ideias pode ser considerada como o núcleo da gestão de inovação e está posicionada no *Front End* da Inovação (BOTHOS; APOSTOLOU; MENTZAS, 2012). Para Sint *et al.* (2010) a Gestão de Ideias é um processo capaz de auxiliar na concepção de novos produtos e no aprimoramento dos processos realizados na organização, podendo ocasionar a minimização dos custos e a manutenção da vantagem competitiva.

De acordo com os autores Aagaard e Gertsen (2011) os investimentos em estudos relativos à Gestão de Ideias são recentes, tanto por parte das organizações, como por parte da academia, mas seus benefícios são reais para as organizações. Benefícios estes que segundo os autores Sint *et al.* (2010); Xie e Zhang (2010); Westerski e Iglesias (2011); Poveda, Westerski e Iglesias (2012), podem acarretar:

- No aumento da produtividade;
- Na melhoria da efetivação dos processos;
- No favorecimento da concepção de novos produtos;
- Na minimização dos investimentos;
- Na manutenção da vantagem competitiva;
- Na criação de valor comercial por meio das ideias identificadas.

A fim de facilitar o gerenciamento de banco de ideias surgiram os sistemas computacionais para o domínio de Gestão de Ideias. Sistemas de Gestão de Ideias tem o objetivo de, através das tecnologias da informação, auxiliar no processo de gerar, avaliar e selecionar ideias inovadoras, utilizando diferentes processos e critérios dentro do mesmo sistema (LI; LI; CHEN, 2014).

Porém, podem ocorrer momentos de sobrecarga de informação para o sistema, devido aos períodos com extensos volumes de submissões de ideias, tornando o processo um desafio para a organização (WESTERSKI; IGLESIAS; RICO, 2010; SINT *et al.*, 2010; WESTERSKI; IGLESIAS, 2011).

Para Magnusson, Netz e Wästlund (2014), um ponto relevante em Sistemas de Gestão de Ideias é o fato de uma mesma ideia aparecer inúmeras vezes na base de dados, embora com especificidades diferentes, porém relacionadas. Isoladamente estas ideias podem não ser interessantes, mas o potencial aumenta quando agrupadas. Ou ainda, caso a mesma ideia seja por vezes recorrente, apontar para uma possível demanda (MAGNUSSON; NETZ; WÄSTLUND; 2014).

Entre as tarefas que podem promover auxílio à Gestão de Ideias encontra-se a análise de agrupamento, ou seja, a descoberta e interpretação de grupos de objetos tendo propriedades e/ou comportamentos semelhantes, sendo uma das operações mais comuns na exploração e análise de vários tipos de dados (ANDRIENKO; ANDRIENKO, 2009). Agrupamentos podem ser detectados em uma meta-análise, indicando ideias semelhantes, apontando padrões e tendências.

A implementação de um módulo em que a informação obtida durante o processo de pesquisa possa ser representada utilizando técnicas de agrupamento, poderá fazer com que usuários interpretem de uma melhor maneira as ideias coletadas na organização (POVEDA; WESTERSKI; IGLESIAS, 2012).

A partir dos elementos identificados esta dissertação objetiva responder a seguinte pergunta de pesquisa: **Como aprimorar o processo de Gestão de Ideias com a aplicação de métodos e técnicas de Engenharia do Conhecimento visando auxiliar na tomada de decisão?**

1.2 OBJETIVOS

Os Objetivos desta dissertação são:

1.2.1 Objetivo Geral

Propor um modelo com foco em métodos e técnicas de Engenharia do Conhecimento capaz de auxiliar na tomada de decisão no contexto da Gestão de Ideias.

1.2.2 Objetivos Específicos

- Propor uma ontologia para o domínio de Gestão de Ideias com foco na tarefa intensiva de análise de agrupamento;
- Propor um modelo que promova suporte ao processo de Gestão de Ideias;
- Demonstrar a viabilidade do modelo proposto por meio da construção de um protótipo, assim como a utilização deste em cenários de estudo;
- Realizar uma discussão dos resultados obtidos a partir dos cenários de estudo.

1.3 JUSTIFICATIVA E RELEVÂNCIA DO TEMA

De acordo com Schreiber *et al.* (2002) por meio da Engenharia do Conhecimento é possível prover métodos para obter entendimento das estruturas e processos utilizados por profissionais que fazem uso do conhecimento de forma intensiva.

Ainda segundo os autores, o conhecimento possui dois aspectos distintos: o senso de finalidade, ao qual está relacionado a um mecanismo intelectual utilizado para atingir metas; e outro relacionado à capacidade de criação, visto que uma de suas principais funções é a de produzir novas informações. Para tanto, atualmente o conhecimento está sendo considerado como um fator de produção.

Newell (1981) considera que o conhecimento pode ser representado como uma combinação de estrutura de dados e procedimentos interpretativos que levam a um comportamento conhecido.

Através da aplicação de técnicas e ferramentas presentes na Engenharia do Conhecimento é possível extrair novos conhecimentos e informações que possam estar presentes no conjunto de dados em análise.

Com o desenvolvimento de sistemas voltados a Gestão de Ideias torna-se possível através das tecnologias da informação auxiliar no processo de gerar, avaliar e selecionar ideias que correspondam a inovações (LI; LI; CHEN, 2014).

Porém, segundo os autores Westerski, Dalamagas e Iglesias (2013), sistemas de Gestão de Ideias geram grandes desafios como:

- O excesso de informação perante o volume de ideias submetidas;
- Os picos súbitos de submissões de ideias;
- A redundância das ideias;
- O volume de ideias triviais;
- O esforço humano dispendido durante o processo de Gestão de Ideias.

O modelo proposto neste trabalho adota técnicas de agrupamento para propiciar a descoberta de grupos de ideias semelhantes. Ideias isoladas podem não representar um potencial de implementação, porém este potencial aumenta quando agrupadas.

Para Poveda, Westerski e Iglesias (2012) a interpretação da informação gerada pela formação dos grupos de ideias pode ser favorecida. A formação dos grupos possui potencial para evidenciar tendências de mercado, e conforme mencionado pelos autores Magnusson, Netz e Wästlund (2014) *clusters* são fatores decisivos no processo de seleção de ideias.

Estudos voltados ao agrupamento de dados atendem a determinados objetivos por propiciarem informações acerca do grau de semelhança ou de diferença entre dois ou mais objetos pertencentes a um conjunto em análise.

Segundo Andrienko e Andrienko (2009) a análise de agrupamento conduz a descoberta e interpretação de grupos de objetos com atribuições e comportamentos semelhantes, sendo uma das operações mais comuns na exploração e análise de diferentes tipos de dados.

Esta técnica possibilita ainda a descoberta não supervisionada de temas e dos principais tópicos relacionados ao conjunto de documentos em análise (ALJABER *et al.*, 2009). A técnica de agrupamento é útil por explorar grandes volumes de dados, uma vez que permite a um analista considerar grupos de objetos em vez de objetos individuais.

O modelo apresenta como forma de estruturação das ideias uma ontologia de domínio objetivando identificar, representar e manipular

este conhecimento, de modo a torná-lo realmente útil para as organizações.

Ontologias proveem a base necessária para o raciocínio sobre os objetos de determinado domínio, devido à representação do conhecimento explicitar o comportamento dos objetos e suas relações (KOBASHI, 2007).

O processo de manutenção da ontologia de domínio é contínuo, pois novas ideias são submetidas ao decorrer do tempo e serão incorporadas à ontologia na forma de novas instâncias.

Para Hüsigg e Kohn (2011) a meta-informação presente na semântica pode, por exemplo, agrupar ideias em categorias, auxiliar na busca e até mesmo avaliação das ideias, extraíndo as ideias mais fortemente discutidas ou as ideias que foram classificadas como as melhores pelo conjunto de usuários.

O valor agregado deste modelo para uma organização está focado essencialmente no auxílio à tomada de decisão e na destinação de recursos para investimentos, relacionados diretamente ao processo de Gestão de Ideias, permitindo que ideias com atribuições semelhantes sejam agrupadas e visualizadas em conjunto. O modelo possibilita também a representação semântica da informação.

1.4 DELIMITAÇÕES DA PESQUISA

O objetivo central deste trabalho é o desenvolvimento de um modelo para auxiliar no processo de tomada de decisão e na seleção de ideias para implementação, estando diretamente relacionado à Engenharia do Conhecimento. Com a formalização deste conhecimento objetiva-se promover um impacto positivo, auxiliando no processo de Gestão de Ideias.

O foco do trabalho centra-se ainda na aplicação das técnicas de indexação e agrupamento para formação dos grupos de ideias. Salienta-se que não faz parte do escopo deste trabalho contribuir ou evoluir os processos acima citados, os quais serão apenas utilizados na concepção do modelo.

Outra característica é a presença de uma ontologia de domínio visando à representação formal e explícita do conhecimento envolvido no processo de Gestão de Ideias. O preenchimento da ontologia e a seleção das características para realização dos agrupamentos é semiautomática, tendo o usuário importante papel na fase de preenchimento e validação das informações.

Esta dissertação baseia-se no pressuposto de que o conhecimento contido nas ideias pode ser melhor visualizado e possivelmente evidenciar padrões e tendências do cenário de mercado por meio das técnicas utilizadas no modelo proposto.

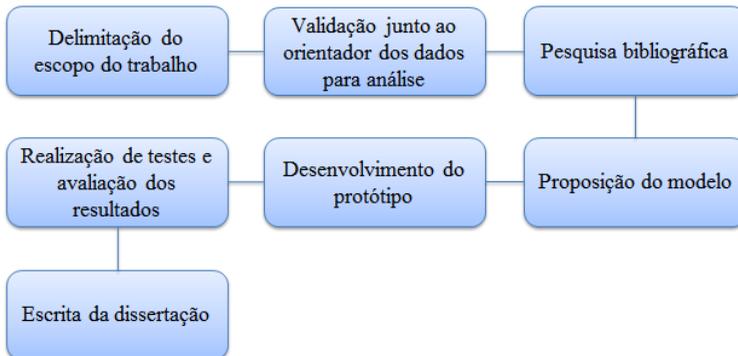
1.5 PROCEDIMENTOS METODOLÓGICOS

Para alcançar os objetivos desta pesquisa, o trabalho decorreu da seguinte forma:

- Realização de uma pesquisa bibliográfica para auxiliar na concepção do modelo proposto e na escrita do trabalho. Os temas abordados na revisão envolveram a área de Gestão de Ideias, Ontologia e *Clustering Analysis* (análise de agrupamento);
- Proposição de um modelo para atender aos objetivos deste trabalho;
- Implementação de um protótipo e aplicação do mesmo em cenários visando auxiliar na demonstração de viabilidade e avaliação do modelo proposto;
- Análise dos resultados obtidos por meio da aplicação do modelo proposto;
- Apresentação das conclusões e perspectivas de trabalhos futuros.

Na Figura 1 são apresentados os procedimentos utilizados para realização desta pesquisa.

Figura 1 - Procedimentos para realização da dissertação



Fonte: Autora.

1.6 ADERÊNCIA E INTERDISCIPLINARIDADE

Esta seção possui como objetivo evidenciar a aderência da dissertação ao objeto de pesquisa do Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento. Para isso, esta seção está dividida em três subseções que apresentam a identidade da dissertação, o contexto estrutural no EGC e as referências factuais.

1.6.1 Identidade

O presente trabalho está relacionado à área de concentração: Engenharia do Conhecimento, na linha de pesquisa: Engenharia do Conhecimento Aplicada às Organizações, cujo foco está em estudar a concepção, desenvolvimento e implantação de soluções de Engenharia do Conhecimento em organizações (EGC, 2014). Esta caracterização se deve ao fato do estudo e da aplicação de técnicas e metodologias que visam melhorar a infraestrutura do conhecimento de uma organização, gerando subsídios que auxiliam a Gestão do conhecimento (GONÇALVES, 2006).

Esta contribuição pode ser reforçada ao analisar-se o objeto de pesquisa do Programa, que está relacionado a macroprocessos de explicitação, gestão e disseminação do conhecimento.

Dentre os macroprocessos estudados pelo Programa, esta dissertação trabalha na gestão, formalização, codificação e explicitação do conhecimento de domínio. A formalização/codificação presente no modelo desenvolvido está baseada em uma ontologia de domínio e na análise de agrupamento, através da formação de grupos de ideias semelhantes. A explicitação do conhecimento ocorre no momento da visualização dos grupos formados, gerando redes e grafos. E a gestão do conhecimento na promoção e suporte ao processo de gerenciamento de ideias, auxiliando na seleção de ideias para implementação e na tomada de decisão.

O objetivo do PPGEGC é investigar, conceber, desenvolver e aplicar modelos, métodos e técnicas relacionados tanto a processos/bens/serviços como ao seu conteúdo técnico-científico (EGC, 2016).

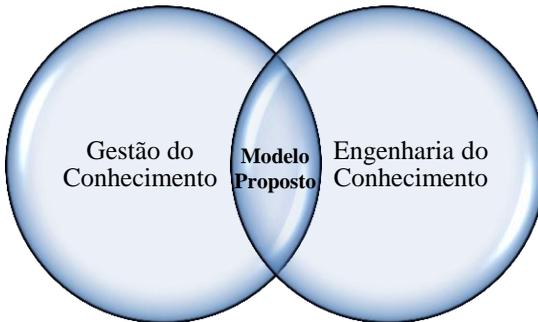
Esta pesquisa constitui caráter interdisciplinar, onde segundo Repko (2008) devem-se avançar as fronteiras disciplinares buscando a criação de novos conhecimentos com base em problemas de pesquisa em comum. Sendo que as atividades atribuídas à área de Engenharia do

Conhecimento relacionam-se com as demais áreas de concentração do programa, Gestão do Conhecimento e Mídia e Conhecimento nos seguintes aspectos:

- Pesquisa e desenvolvimento de metodologias de identificação, representação e gestão de conhecimento;
- Aplicação de sistemas de conhecimento à gestão do conhecimento organizacional (formalização, memória e tomada de decisão);
- Aplicação de sistemas de conhecimento à interação homem-máquina, como suporte aos trabalhadores de conhecimento, inclusive na educação;
- Aplicação de sistemas de conhecimento em mundos virtuais interativos visando à melhoria da eficácia e eficiência dos processos de treinamento e capacitação (EGC, 2016).

A Figura 2 retrata a integração das áreas de concentração do Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento envolvidas na pesquisa:

Figura 2 - Áreas de concentração envolvidas na pesquisa



Fonte: Autora.

1.6.2 Contexto estrutural no EGC

A área de Engenharia de Conhecimento surgiu a partir da Inteligência Artificial. O principal produto da área era os sistemas baseados em conhecimento. A partir dos anos 90 ocorreu a

intensificação do papel do conhecimento como elemento estratégico nas organizações.

Devido à popularidade dos sistemas baseados em conhecimento surgiu a necessidade de uma abordagem sistemática para a construção de tais sistemas, similares aos procedimentos adotados na Engenharia de *Software* (STUDER et al., 1998).

Para Studer, Benjamins e Fensel (1998) a área de Engenharia do Conhecimento envolve a aplicação de teorias, métodos e ferramentas voltadas à modelagem e representação do conhecimento de domínio (STUDER, BENJAMINS; FENSEL, 1998), provendo um conjunto de ferramentas para formalizar e explicitar este conhecimento, de modo a auxiliar nas tarefas intensivas em conhecimento e assim fornecer elementos para a Gestão do Conhecimento (SCHREIBER et al., 2002).

Ainda segundo Schreiber (2007), ela fornece uma orientação sobre qual o momento e os procedimentos para aplicar técnicas de representação de conhecimento para resolver problemas específicos.

Para tal, os objetivos da área de Engenharia do Conhecimento incluem:

A pesquisa e o desenvolvimento de técnicas e ferramentas para a formalização, codificação e gestão do conhecimento; métodos de análise da estrutura e processos conduzidos por profissionais em atividades de conhecimento intensivo; e a pesquisa e desenvolvimento de sistemas de conhecimento (EGC, 2016).

No que tange ao escopo desta dissertação, o aspecto que a contextualiza na área de Engenharia do Conhecimento reside no fato do modelo possuir como objetivo a materialização, principalmente dos macroprocessos de explicitação, formalização/codificação e gestão do conhecimento.

Ao desenvolver um sistema computacional voltado ao agrupamento de ideias, foco deste trabalho, especialistas de domínio podem interpretar e visualizar melhor as ideias coletadas na organização (POVEDA; WESTERSKI; IGLESIAS, 2012). Segundo os autores Hüsigg e Kohn (2011) através da semântica presente na formalização/Codificação deste conhecimento por meio de uma ontologia voltada ao domínio de Gestão Ideias, é possível agrupar ideias em categorias, auxiliar na busca e até mesmo na avaliação das ideias, extraindo as ideias mais fortemente discutidas ou as ideias que foram classificadas como as melhores pelo conjunto de usuários. E por meio de

uma meta-análise evidenciar grupos de ideias semelhantes que podem indicar possíveis padrões, tendências e demandas de mercado (MAGNUSSON; NETZ; WÄSTLUND, 2014).

1.6.3 Referências Factuais

O presente trabalho possui como foco principal as áreas de Representação Semântica, Gestão de Ideias e Análise de Agrupamento. A seguir, serão apresentados os trabalhos desenvolvidos no programa que estão relacionados ao foco do trabalho.

Na área de Representação Semântica, foram encontrados 33 trabalhos, sendo os mais atuais:

- BOTELHO, Maurício. Aplicação de Ontologias na Organização de Conteúdos para Apoio a Equipes de Desenvolvimento de *Software*. Dissertação, 2015.
- CECI, Flavio. Um Modelo Baseado em Casos e Ontologia para Apoio a Tarefa Intensiva em Conhecimento de Classificação com Foco na Análise de Sentimento. Tese, 2015.
- SILVA, Thales do Nascimento da. Um Modelo Baseado em Ontologia para Suporte a Tarefa Intensiva em Conhecimento de Recomendação. Dissertação, 2015.
- BRAGLIA, Israel. Um Modelo Baseado em Ontologia e Extração de Informação como Suporte ao Design Instrucional na Geração de Mídias do Conhecimento. Tese, 2014.
- CARDENAS, Yuri Gomes. Modelo de Ontologia para Representação de Jogos Digitais de Disseminação do Conhecimento. Dissertação, 2014.

No domínio de Gestão de Ideias foram encontrados poucos trabalhos, sendo os mais relacionados:

- DOROW, Patrícia Fernanda. O Processo de Geração de Ideias para Inovação: Estudo de Caso em uma Empresa Náutica. Dissertação, 2013.
- PRADA, Charles A. Proposta de modelo para o gerenciamento de portfólio de inovação: modelagem do conhecimento na geração de ideias. Dissertação, 2009.

- MIGUEZ, Viviane Brandão. Uma Abordagem de Geração de Ideias para o Processo de Inovação. Dissertação, 2012.

E na área de Análise de Agrupamento o trabalho relacionado foi o de:

- CECI, Flávio. Um Modelo Semi-automático Para a Construção e Manutenção de Ontologias a partir de bases de documentos não estruturados. Dissertação, 2010.

Diante das referências apresentadas, a presente dissertação é aderente ao Programa por propor um modelo de conhecimento. Compreende-se que o presente trabalho está de acordo com a área de concentração de Engenharia do Conhecimento, objetivando dar suporte a área de Gestão do Conhecimento, e possui trabalhos anteriores que abordam temáticas semelhantes.

1.7 ESTRUTURA DO TRABALHO

O presente trabalho é composto de cinco capítulos descritos a seguir:

- O primeiro capítulo apresenta o tema e os objetivos desta pesquisa, suas delimitações e aderência ao programa de pós-graduação;
- O segundo capítulo é composto de um referencial teórico, no qual são abordados os assuntos principais relacionados à pesquisa. Os assuntos são: Gestão de Ideias, Ontologia e Análise de Agrupamento;
- No capítulo 3, apresenta-se o modelo proposto por meio de uma descrição detalhada sobre as etapas que o compõem, sendo: (1) Indexação; (2) Validação; (3) Agrupamento; (4) Visualização;
- No capítulo 4, é apresentada a proposição de avaliação do modelo através da discussão dos resultados alcançados por meio dos cenários de estudo;
- O quinto e último capítulo apresenta as conclusões da dissertação e as sugestões de trabalhos futuros. Por fim, são disponibilizados as referências utilizadas na pesquisa e os apêndices.

2 FUNDAMENTAÇÃO TEÓRICA

A fundamentação teórica apresenta os principais conceitos relacionados ao desenvolvimento do modelo voltado a Gestão de Ideias, baseado em Ontologia e na Análise de Agrupamento. Neste capítulo, abordam-se as definições de Gestão de Ideias, Sistemas de Gestão de Ideias, Ontologia e Análise de Agrupamento. Apresenta-se o processo de agrupamento de ideias, uma descrição sobre o processo de mineração de texto para descoberta de conhecimento e uma explanação sobre os algoritmos hierárquicos e dos diferentes algoritmos por particionamento.

2.1 GESTÃO DE IDEIAS

A competitividade empresarial fez com que as organizações buscassem investir em novos produtos e estratégias de atuação. Em face da competitividade, a inovação tornou-se um fator crucial para sustentar a organização no mercado (GIBSON; SKARZYNSKI, 2008), o que representa um mecanismo complexo capaz de transformar processos em vantagens competitivas (BAUTZER, 2009).

Marzano (2005) define inovação como um processo sistemático para a criação e incremento da vantagem competitiva. Para Clark e Wheelwright (1993) a inovação está diretamente relacionada à sustentabilidade das organizações que coexistem em um cenário competitivo e necessitam conquistar novos mercados. Segundo Moos *et al.* (2011) a inovação é uma das funções fundamentais para uma empresa e uma necessidade competitiva em um mercado dinâmico.

Bautzer (2009) afirma que por meio da inovação é possível agregar valor aos produtos e serviços, contudo, compreender a capacidade em gerar inovações por parte das organizações é fundamental (QUINTANE, 2011). Para Le (2004) organizações que trabalham com projetos inovadores possuem maior habilidade em gerir mudanças e conseguem mitigar de maneira mais adequada os riscos inerentes ao processo.

Ainda segundo os autores Schumpeter (1949), Gaynor (2001), Adams; Bessant; Phelps (2006), Jensen; Webster (2009), Oke; Munshi; Walumbwa (2009), o termo inovação está associado à geração de lucros, a valor de mercado e à exploração bem sucedida de ideias.

Baregheh, Rowley e Sambrook (2009) definem inovação como a habilidade da organização em transformar ideias em produtos novos ou melhorados, serviços ou processos, buscando o crescimento, a

capacidade de competir e a busca de diferenciais capazes de garantir o sucesso de mercado.

Para os autores Gambatese e Hallowell (2011) o termo inovação está relacionado a três componentes, sendo estes, à geração de oportunidades, a difusão e a geração de ideias. Inovações surgem por intermédio de transformações positivas introduzidas por novas ideias (GAMBATESE; HALLOWELL, 2011).

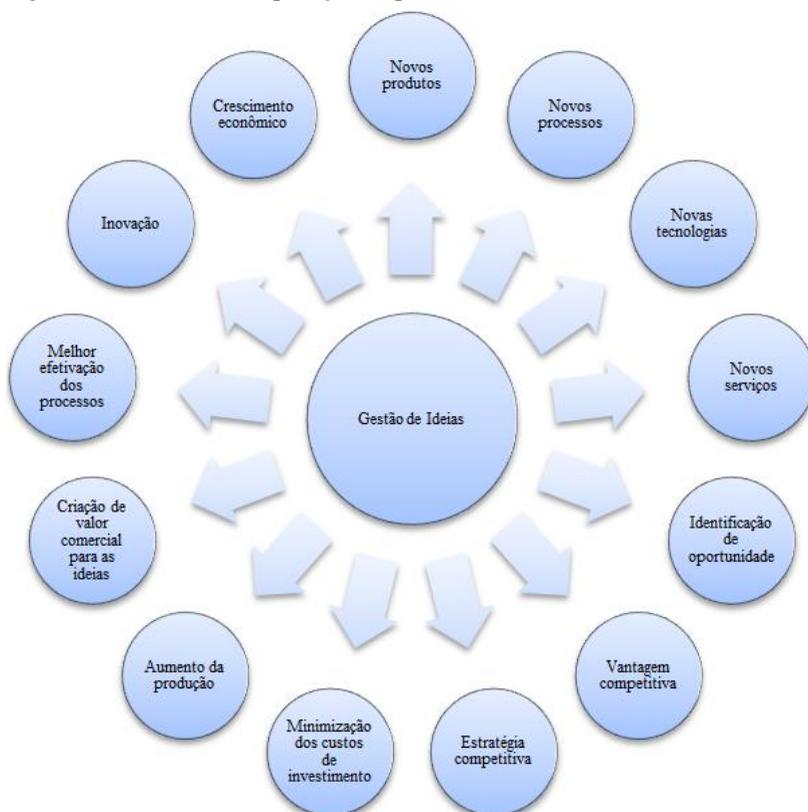
Para Art *et al.* (2010) os investimentos em inovação são necessários devido ao aumento do nível de concorrência e diminuição do ciclo de vida do produto. Ainda segundo os autores, através destes investimentos é possível aumentar a rentabilidade e nutrir a vantagem competitiva organizacional.

Mahroum e Al-Saleh (2013) afirmam que em uma economia globalizada a capacidade de gerar inovações e ideias e colocá-las em prática tornou-se uma estratégia poderosa para o crescimento econômico. Para Flynn *et al.* (2003) a capacidade de crescimento da organização depende da competência em gerar novas ideias e explorá-las efetivamente para o benefício da organização a longo prazo. Para Endesley (2010) “o coração da inovação é a geração e o teste de novas ideias”.

Barbieri, Álvares e Cajazeira (2009) salientam que a Gestão de Ideias é parte fundamental no processo de inovação. Flynn *et al.* (2003) também corroboram desta afirmação ao alegarem que a Gestão de Ideias proporciona o desenvolvimento de inovações, concretizando-se como uma etapa essencial do processo de inovação (COOPER; EDGETT; KLEINSCHMIDT, 2001).

Com base nos trabalhos de Poveda, Westerski e Iglesias (2012); Westerski e Iglesias (2011); Witell, Löfgren e Gustafsson (2011); Björk, Boccardelli e Magnusson (2010); Sint *et al.* (2010); Xie e Zhang (2010); Baregheh, Rowley e Sambrook (2009); Brem e Voigt (2007); Vandenbosch, Saatcioglu e Fay (2006); Bessant *et al.* (2005) e Flynn *et al.* (2003); em decorrência das citações referentes à Gestão de Ideias, a Figura 3 representa os benefícios obtidos com a aplicação do processo.

Figura 3 – Benefícios da aplicação do processo de Gestão de Ideias



Fonte: Autora.

Ideias representam a matéria-prima para o processo de criação de novos produtos (COOPER; EDGETT; KLEINSCHMIDT, 2002). Segundo Iqbal (2011) gerar ideias é um processo conexo à criatividade intraorganizacional. Segundo o autor a inovação concretiza-se com o desenvolvimento dessas ideias. Para Sandström e Björk (2010) a Gestão de Ideias possui como objetivo originar, avaliar e eleger ideias para inovação.

Adams, Bessant e Phelps (2006) salientam que os pontos críticos do processo estão associados à filtragem e seleção das ideias, sendo estes, fatores impactantes na falha ou sucesso da organização.

Os autores Westerski e Iglesias (2011) também corroboram desta afirmação, alegando que dentre as etapas do processo de Gestão de

Ideias a mais importante e problemática está relacionada à avaliação das ideias.

Identificar as ideias inovadoras não é um processo trivial, uma vez que estas se encontram tanto dentro, quanto fora da organização (COOPER; EDGETT; KLEINSCHMIDT, 2002; MCGRATH, 2004). O autor VanGundy (2007) destaca a importância de alinhar e especificar o alvo, a direção e as prioridades da empresa para, num processo subsequente, promover a Gestão de Ideias.

Contudo, os esforços e estudos por parte da academia e das organizações para promoção da Gestão de Ideias são recentes (AAGAARD; GERTSEN, 2011). Xie e Zhang (2010) também confirmam esta afirmação, salientando que poucas são as pesquisas voltadas ao gerenciamento do processo de geração, seleção, melhoria e implementação de ideias.

Os autores Mowell e Shea (2001) evidenciam a dificuldade enfrentada pelas organizações em identificar as ideias inovadoras com real potencial para implementação, capazes de gerar resultados rentáveis para a organização.

2.1.1 Sistemas de Gestão de Ideias

Os Sistemas de Gestão de Ideias, do inglês, *System Management Idea* (IMS) surgem com o objetivo de gerenciar as ideias coletadas na organização. Permitem que as empresas gerenciem o processo de descoberta, incubação, aceleração e dimensionamento das ideias para produzir valor comercial por meio do desenvolvimento de produtos e processos inovadores (POVEDA; WESTERSKI; IGLESIAS, 2012).

Estes sistemas fornecem ferramentas que auxiliam na avaliação e seleção de ideias para implementação (WESTERSKI; IGLESIAS; RICO, 2010). Os autores Sandstrom e Bjork (2010) também corroboram desta afirmação e salientam que os sistemas de Gestão de Ideias visam produzir, avaliar e selecionar ideias de inovação contínua e descontínua.

Para Westerski e Iglesias (2011) os sistemas computacionais voltados a Gestão de Ideias proporcionam a organização e a avaliação das ideias, com o intuito de gerar um conjunto com as ideias que possuem potencial real de implementação e valor para a organização. Ferramentas e técnicas associadas ao processo de Gestão de Ideias aumentam substancialmente a qualidade e o número de ideias (MCADAM; MCCLELLAND, 2002).

Estes sistemas trabalham com a coleta e organização das ideias relacionadas às sugestões de inovações de produtos e serviços, geradas pelos integrantes da organização, concretizando como um ramo promissor da indústria (PETTEY; STEVENS, 2009).

Segundo Conry-Murray (2010) o interesse crescente em sistemas de Gestão de Ideias tem impulsionado uma comunidade no desenvolvimento de ferramentas, com a finalidade de agregar o processo de gerenciamento de ideias ao processo de desenvolvimento de *software*.

Sistemas que trabalham com a Gestão de Ideias utilizam o conceito de Inteligência Coletiva (IC), que só se tornou viável após o desenvolvimento da Web 2.0 e suas ferramentas de interação (CONVERTINO; GRASSO, MICHELIS, 2010). Por meio das tecnologias da informação e comunicação (TICs), IMS podem ser utilizados para monitorar o processo de Gestão de Ideias, auxiliando na captura e triagem das ideias (COOPER; EDGETT, 2009).

Organizações como a Dell®, Starbucks®, Cisco® e a Canonical® vem desenvolvendo aplicativos de Gestão de Ideias com a intenção de impulsionar o envolvimento dos usuários para estimular a inovação e melhorar seus produtos e serviços (WESTERSKI; DALAMAGAS; IGLESIAS, 2013).

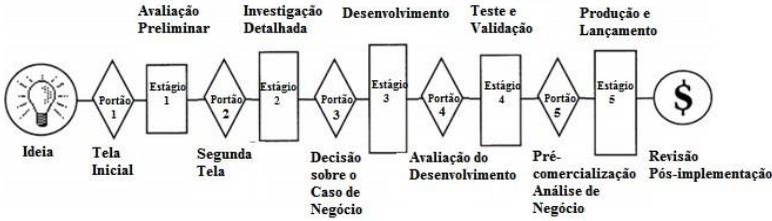
2.1.2 Modelos para Gestão de Ideias

Os modelos presentes no domínio de Gestão de Ideias objetivam prover suporte ao processo de gerenciamento de ideias.

A inovação pode ser vista de duas formas, como processo, ou como resultado. No modelo sugerido por Cooper (1990), a inovação é apresentada como um processo. De acordo com o autor, como todo processo, este pode ser gerenciado. Ao longo do tempo o modelo sofreu atualizações, porém a essência não foi alterada.

Segundo Cooper (1990) a ideia é o primeiro passo para o processo de desenvolvimento de um produto. O modelo é composto por várias fases. A primeira fase corresponde à avaliação preliminar do projeto. A segunda é referente à definição sobre a implementação do projeto. A terceira ao desenvolvimento. Posteriormente são realizados os testes e a validação do projeto, e a quinta refere-se à comercialização. O modelo apresenta ainda como etapa final a revisão pós-implementação. A Figura 4 apresenta o modelo.

Figura 4 – Modelo *Stage Gate*

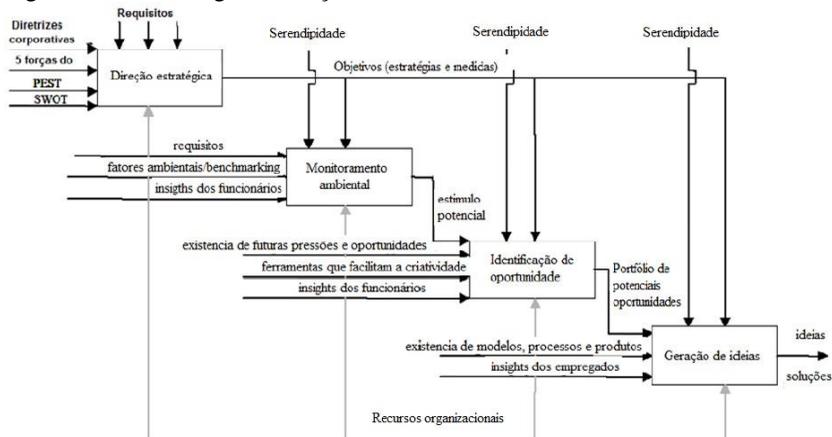


Fonte: Cooper (1990)

Os autores Flynn *et al.* (2003) expõem uma metodologia fundamentada nos colaboradores da organização, nas habilidades, nos conhecimentos e na comunicação, composta por quatro etapas: a direção estratégica, o monitoramento ambiental, a identificação de oportunidades e a etapa de geração de ideias. O resultado de cada etapa presente na metodologia atua como o princípio para a próxima, podendo ainda influenciar nas demais (FLYNN *et al.*, 2003).

A metodologia foi criada para promoção da criatividade em um ambiente de negócios. Nesta metodologia é possível prover suporte ao processo de criação de ideias, ao alinhamento das metas e ao reconhecimento da oportunidade de definição da ideia, com o intuito de alinhar a ideia à estratégia da organização e a identificação de oportunidade, para a escolha da ideia a ser colocada em prática. A metodologia é apresentada na Figura 5.

Figura 5 – Metodologia de criação de ideias



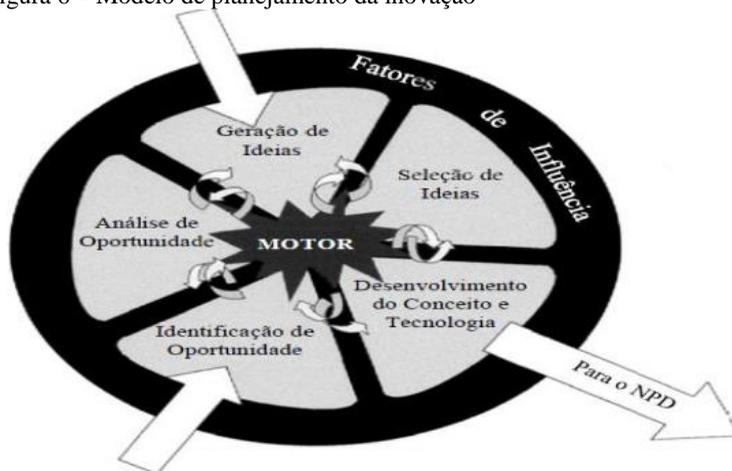
Fonte: Flynn *et al.* (2003, p. 427)

O modelo apresentado por Koen *et al.* (2002) expõe um formato interativo e não sequencial entre os elementos de acordo com o movimento das setas. Os autores criaram o modelo a fim de subsidiar as atividades elaboradas antes do desenvolvimento formal dos produtos.

O modelo trata do processo de inovação como algo complexo, de difícil planejamento, composto por imprevisibilidades. O modelo foi nomeado de: Modelo de Desenvolvimento de Novo Conceito (do inglês, *New Concept Development Model*). Os autores consideram a existência de cinco atividades, sendo elas: a identificação de oportunidades, a análise de oportunidade, a geração de ideias, a seleção de ideias e a definição de conceito.

No centro do modelo encontra-se o motor, apresentado na Figura 6, ao qual é responsável por coordenar as cinco atividades apresentadas. Possui ainda fatores externos que influenciam no processo, como por exemplo, a concorrência de mercado.

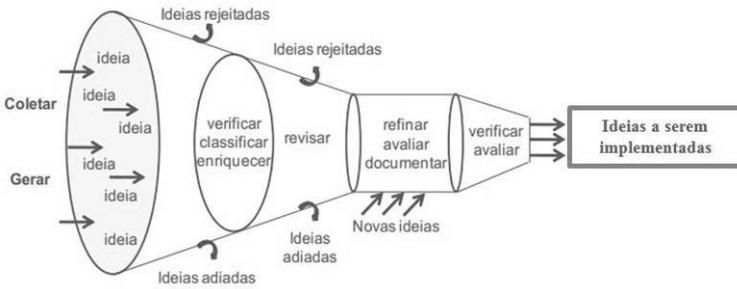
Figura 6 – Modelo de planejamento da inovação



Fonte: Koen *et al.* (2002)

Brem e Voight (2007) em seu modelo, enfatizaram a importância da implementação da Gestão de Ideias nas organizações para obtenção de inovações positivas. O modelo decompõe o processo de inovação em três fases: a geração de ideias, a aceitação da ideia e a concretização da ideia. Os autores desenvolveram o túnel de ideias baseado no funil de desenvolvimento, conforme apresentado na Figura 7.

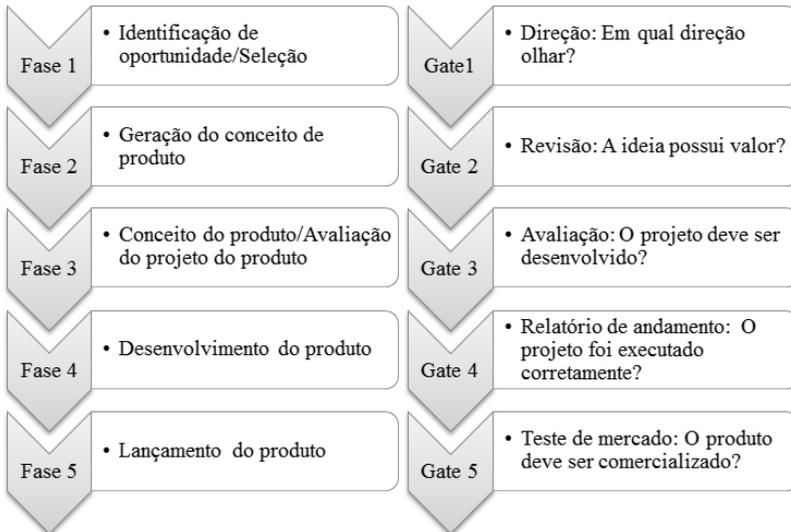
Figura 7 – Túnel de ideias



Fonte: Deschamps, Nayak e Little (1996)

O modelo proposto por Crawford e Benedetto (2006), inicia com a identificação de uma ideia e termina com a aprovação de um projeto para execução. Os autores propuseram um modelo para desenvolvimento de um processo de inovação em produtos, onde acreditam que através da combinação de passos, atividades, objetivos e decisões, quando executadas corretamente, geram inovações para a organização. A Figura 8 apresenta o modelo desenvolvido por Crawford e Benedetto.

Figura 8 – Processo de inovação segundo Crawford e Benedetto

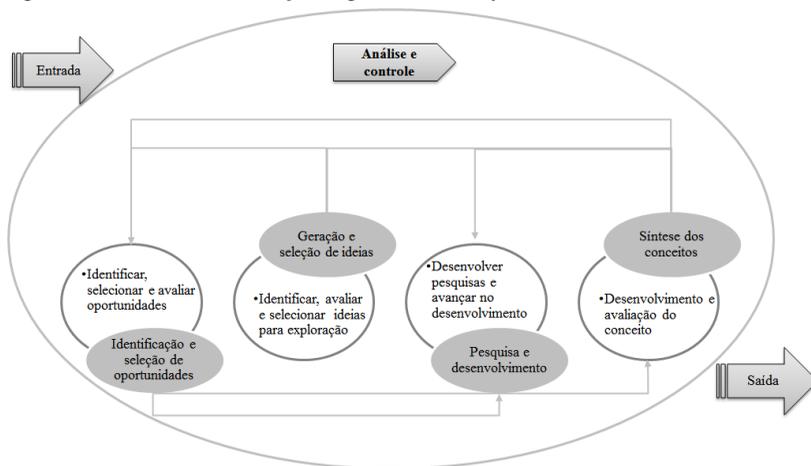


Fonte: Adaptado de Crawford e Benedetto (2006)

A fase 1 do modelo busca identificar oportunidades de negócio. A fase 2 objetiva levantar novas ideias de acordo com as oportunidades identificadas na primeira fase. A fase 3 corresponde à avaliação do projeto e envolve três atividades principais: a análise e a seleção das ideias para implementação, a avaliação técnica e comercial da proposta e a definição do projeto. A fase 4 corresponde à execução do projeto e na fase 5 são realizados testes no produto para posterior comercialização.

Whitney (2007) propõe um modelo para o desenvolvimento de tecnologias. Por meio de um conjunto de ferramentas utilizadas no processo de inovação, torna-se o processo iterativo. O modelo possui ainda um retorno das avaliações, o que ajuda a equilibrar e gerenciar o processo. A Figura 9 apresenta o modelo.

Figura 9 – Modelo de inovação segundo Whitney



Fonte: Adaptado de Whitney (2007)

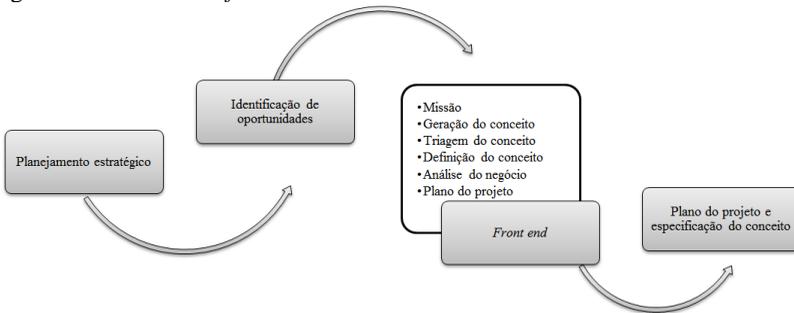
No modelo de Whitney as entradas para o sistema são atividades que estimulam o processo de inovação. Já as saídas são a melhoria das atividades ou ainda o desenvolvimento de uma nova tecnologia. Os passos que estas atividades ou ideias irão percorrer são primeiramente: a identificação de oportunidades, cujo objetivo é identificar, selecionar e avaliar oportunidades. Em seguida, ocorrerá a geração e seleção de ideias com o intuito de identificar, avaliar e selecionar ideias a serem exploradas. O terceiro passo corresponde à pesquisa e desenvolvimento (diferencial deste modelo, pois enfatiza a área de pesquisa) para

desenvolver, pesquisar e avançar no desenvolvimento da ideia ou atividade. Por último, ocorre à síntese dos conceitos, objetivando o desenvolvimento e avaliação do conceito criado a partir da ideia ou atividade que poderá originar a melhoria ou desenvolvimento de uma tecnologia.

Nobelius e Trygg em 2002 propuseram um modelo para o *front end* do processo de inovação. Primeiramente é necessário definir um conceito antes que o processo de desenvolvimento do projeto inicie efetivamente. Os autores enfatizam a necessidade de execução de todas as fases do modelo para o sucesso da atividade.

Nobelius e Trygg (2002) destacam que todos os modelos existentes relacionados ao *front end* possuem um planejamento estratégico. A saída destes modelos geram um plano de projeto e a especificação do conceito do produto. Ao analisar os modelos existentes, os autores fizeram uma proposição baseada na síntese dos demais e concluíram que os processos essenciais no *front end* consistiam na declaração da missão, criação do conceito, triagem/definição do conceito, análise do negócio e, por fim, planejamento do projeto. A Figura 10 apresenta o modelo proposto por Nobelius e Trygg.

Figura 10 – Modelo de *front end*



Fonte: Adaptado de Nobelius e Trygg (2002)

2.2 ONTOLOGIA

O termo ontologia origina-se na Filosofia e está relacionado à existência do ser ou aos tipos de existência (BORST, 1997). Na inteligência artificial o termo está associado ao raciocínio exercido por um sistema de IA para execução de determinada tarefa.

Segundo os autores Poli e Obrst (2010), atualmente o termo ontologia por ser visualizado de acordo com duas perspectivas: a primeira está relacionada à filosofia, mencionada no primeiro parágrafo. A segunda está relacionada à Ciência da Computação, inicialmente utilizada pela Inteligência Artificial, e nos dias atuais, utilizada também pela área de Engenharia do Conhecimento.

A pesquisa desenvolvida utiliza o conceito de ontologia baseada na visão da Ciência da Computação e da Engenharia do Conhecimento, onde a mesma é vista como um modelo de dados para modelagem e representação do conhecimento de domínio, possibilitando realizar inferências sobre os objetos do domínio.

Nesta perspectiva, em 1993, Gruber atribuiu significado ao termo, definindo-o como a exploração e representação dos relacionamentos semânticos e conceituais da informação. Guarino em 1998 definiu o termo como uma especificação das propriedades e relações entre os conceitos de um determinado domínio.

Para Guarino (1998) o termo está associado à definição de um vocabulário para descrever determinada realidade. Neches *et al.* (1991) vão além e afirmam que por meio da ontologia é possível determinar as regras para combinar termos e relacionamentos, definindo extensões do vocabulário.

Uma das definições mais aceitas para o termo foi fundamentada em Gruber (1993) e Borst (1997), onde os autores Studer, Benjamins e Fensel (1998) afirmaram que: “a ontologia é a representação formal e explícita de uma conceitualização compartilhada”. Para Vasconcelos, Rocha e Kimble (2003) por meio da ontologia é possível especificar formalmente e em alto nível um domínio.

Ontologias são utilizadas como uma forma de comunicação e integração do conhecimento, bem como, para realizar inferência sobre este conhecimento (MIKA; AKERMANS, 2005). Promovem a representação do conhecimento através de termos, definindo conceitos do mundo real, referentes a um determinado domínio, representando as relações semânticas e não somente as relações sintáticas entre os dados.

Ao desenvolver uma ontologia, componentes como, a descrição das categorias e dos objetos e as relações entre os dados envolvidos no processo são explicitados (LULA; PALIWODA-PEKOSZ, 2008).

Para Suárez-Figueroa *et al.* (2011) classes, relacionamento, axiomas e instâncias formam o conjunto de componentes que compõem a representação do conhecimento de domínio. Maedche (2002) afirma que a descrição de uma ontologia deve ser composta por 5-tuplas sendo:

os conceitos, os relacionamentos, a hierarquia de conceitos, funções que relacionam conceitos e um conjunto de axiomas.

Para Uschold e Jasper (1999) ontologias assumem diferentes formatos, incluindo fundamentalmente um vocabulário de termos e sua especificação, as definições dos relacionamentos entre os conceitos, resultando na estruturação do domínio e suas restrições.

Ontologias são instrumentos com alto poder de organização e representação da informação de domínio, otimizando os processos de recuperação de informação (ALMEIDA; BAX, 2003). Estas podem ser divididas de duas formas, de acordo com o grau de formalismo podendo representar apenas uma taxonomia, ou um vocabulário controlado (*light-weight*), ou ainda especificar os relacionamentos e restrições lógicas entre os conceitos (*heavy-weight*) (BATZIAS, SIONTOROU, 2012).

Para Kobashi (2007) a ontologia fornece a base necessária para o raciocínio sobre os objetos de determinado domínio, pois a partir da representação do conhecimento ocorre a promoção da explicitação do comportamento dos objetos e relações. Gómez-Pérez (1999) afirma que uma ontologia não inclui somente os termos que podem ser explicitados, mas também os que podem ser obtidos por meio de regras de inferência.

O uso desta tecnologia de representação permite a habilidade em raciocinar, analisar e operar sobre o conhecimento armazenado. Segundo Varshovi e Sadeghiyan (2010) sua utilização facilita o processo de raciocínio e cooperação de sistemas de detecção e resposta.

A meta-informação presente na semântica pode, por exemplo, agrupar ideias em categorias, auxiliar na busca e até mesmo na avaliação das ideias, extraindo as ideias mais fortemente discutidas ou as ideias que foram classificadas como as melhores pelo conjunto de usuários (HÜSIG; KOHN, 2011).

Guarino (1998) propôs a seguinte definição para os diferentes tipos de ontologia:

- Ontologias gerais (*top-level ontology*): possuem significados abstratos para a compreensão do domínio de conhecimento;
- Ontologias de domínio (*domain ontology*): abordam um domínio específico de uma área genérica;
- Ontologias de tarefa (*task ontology*): abordam tarefas ou atividades genéricas;
- Ontologias de aplicação (*application ontology*): objetivam solucionar um problema específico de um determinado domínio, normalmente referenciando termos relacionados a uma ontologia de domínio.

Studer (1998) considera também outros dois tipos de ontologias:

- Ontologias de representação: definem conceitos que especificam genericamente a representação do conhecimento, não se detendo a um domínio específico;
- Ontologias de método: especificam o vocabulário relativo a um método presente em um Método de Resolução de Problema (PSM).

Almeida e Bax (2003) em seus estudos propuseram uma especificação ainda mais completa sobre os tipos de ontologia, apresentada no Quadro 1.

Quadro 1 - Tipos de ontologias

Abordagem	Classificação	Definição
Quanto à função Mizoguchi, Vanwelkenhuysen e Ikeda (1995)	Ontologias de domínio	Reutilizáveis no domínio fornecem vocabulário sobre conceitos, seus relacionamentos, sobre atividades e regras que os governam.
	Ontologias de tarefa	Fornecem um vocabulário sistematizado de termos, especificando tarefas que podem ou não estar no mesmo domínio.
	Ontologias gerais	Incluem um vocabulário relacionado a coisas, eventos, tempo, espaço, casualidade, comportamento, funções, etc.
Quanto ao grau de formalismo Uschold e Gruninger (1996)	Ontologias altamente informais	Expressa livremente em linguagem natural.
	Ontologias semi-informais	Expressa em linguagem natural de forma restrita e estruturada.
	Ontologias semiformais	Expressa em linguagem artificial definida formalmente.
	Ontologia rigorosamente formal	Os termos são definidos com semântica formal, teoremas e provas.
Quanto à aplicação Jasper e Uschold (1999)	Ontologias de autoria neutra	Um aplicativo é escrito em uma única língua e depois convertido para uso em diversos sistemas, reutilizando-se as informações.
	Ontologias como especificação	Cria-se uma ontologia para um domínio, a qual é usada para documentação e manutenção no desenvolvimento de softwares.
	Ontologia de acesso comum à informação	Quanto ao vocabulário é inacessível, a ontologia torna a informação inteligível, proporcionando vocabulário compartilhado dos termos.
Quanto à estrutura	Ontologia de alto-	Descrevem conceitos gerais

Ontologia de alto-nível Haav e Lubi (2001)	nível	relacionados a todos os elementos da ontologia (espaço, tempo, matéria, objeto, evento, ação, etc.) os quais são independentes do problema ou domínio.
	Ontologia de domínio	Descrevem o vocabulário relacionado ao domínio, como, por exemplo, medicina, ou automóveis.
	Ontologia de tarefa	Descrevem uma tarefa ou atividade, como, por exemplo, diagnósticos ou compras, mediante inserção de termos especializados na ontologia.
Quanto ao conteúdo VanHeijst, Schreiber e Wielinga (2002)	Ontologias terminológicas	Especificam termos que serão usados para representar o conhecimento em um domínio (por exemplo, os léxicos).
	Ontologias de informação	Especificam a estrutura de registros de bancos de dados (por exemplo, os esquemas de bancos de dados).
	Ontologias de modelagem do conhecimento	Especificam conceituações do conhecimento, tem uma estrutura interna semanticamente rica e são refinadas para uso no domínio do conhecimento que descrevem.
	Ontologias de aplicação	Contém as definições necessárias para modelar o conhecimento em uma aplicação.
	Ontologias de domínio	Expressam conceituações que são específicas para um determinado domínio do conhecimento.
	Ontologias genéricas	Similares às ontologias de domínio, mas os conceitos que as definem são considerados genéricos e comuns a vários campos.
	Ontologias de representação	Explicam as conceituações que estão por trás dos formalismos de representação do conhecimento.

Fonte: Almeida e Bax (2003, p. 10)

A utilização de ontologia atualmente tornou-se frequente para possibilitar a representação do conhecimento e apoiar tecnologias voltadas à gestão do conhecimento (FERNANDES *et al.*, 2011). Segundo os autores Zanni-merk; Cavallucci e Rousselot (2009), Carbone *et al.* (2012), Kumar *et al.* (2010), Gardner (2005) e Jung (2009), ontologias auxiliam no processo de:

- Aquisição do conhecimento;
- Criação do conhecimento;
- Compartilhamento do conhecimento;

- Distribuição do conhecimento;
- Integração entre sistemas;
- Interoperabilidade entre sistemas.

2.2.1 Trabalhos correlatos relacionados ao domínio de Gestão de Ideias e Aplicações Semântica

Esta seção consiste na identificação das principais características do objeto de estudo, a partir de um conjunto de artigos selecionados e analisados no ano de 2016, através das principais bases de dados: *Scopus*, *ACM*, *IEEE* e *Springer Link*.

Com o intuito de analisar as aplicações semânticas no cenário de gestão de ideias, foi utilizada a seguinte expressão como filtro de pesquisa: (“*semantic*” and “*idea management*”) e também utilizou-se a expressão (“*ontology*” and “*idea management*”) nos campos referentes a título, palavras-chave e resumo, nas bases que dispunham deste recurso. A pesquisa pelo termo ontologia foi necessária devido à expressividade e utilização do termo nas pesquisas semânticas. Foram identificados 21 artigos, no entanto somente 6 apresentavam uma solução semântica para o domínio de Gestão de Ideias.

No artigo denominado de *Idea management system for team creation (2010)*, é realizada uma proposição de um sistema capaz de gerenciar o processo de Gestão de Ideias, promovendo a criação de equipes. Como resultado os autores desenvolveram um IMS (sistema de gestão de ideias, do inglês, *Idea Management System*), capaz de realizar o reconhecimento de uma ideia criativa por meio de análise semântica latente, da seleção de uma ideia e da evolução e visualização da mesma (XIE; ZHANG, 2010).

No artigo intitulado de *Ideator - a collaborative enterprise idea Management tool powered by kiwi (2010)*, ocorre o desenvolvimento de uma ferramenta para apoiar o processo de gestão de ideias. O Ideator é uma aplicação Enterprise 2.0, capaz de promover suporte a geração de ideias de maneira colaborativa baseada na semântica promovida pelo *framework KiWi* (SINT et al., 2010).

No trabalho *A model for integration and interlinking of idea Management systems (2010)*, ocorre a introdução da utilização de tecnologias da Web Semântica entre Sistemas de Gestão de Ideias e *softwares* heterogêneos, capaz de alcançar a interoperabilidade entre os mesmos. O modelo propõe como e quais tipos de anotações de metadados devem ser aplicados no domínio dos sistemas de Gestão de

Ideias. A Web Semântica pode funcionar como ferramenta para criar novas oportunidades e alavancar os sistemas atuais de Gestão de Ideias (WESTERSKI; IGLESIAS; RICO, 2010).

No artigo denominado de *Exploiting structured linked data in enterprise knowledge management systems: An idea management case study (2011)*, existe a proposição de um modelo de criação de dados abertos ligados para a *World Wide Web*. Como resultados alcançados, existe a promoção da interligação de dados estruturados para uso em sistemas voltados a Gestão de Ideias, em face da quase inexistência de sistemas de Gestão do Conhecimento dedicados à inovação. E ainda, o desenvolvimento de um mecanismo de extensão de uma ontologia que primeiramente abrangia apenas a estrutura dos sistemas voltados à Gestão de Ideias, para o conceito de vinculação com dados corporativos distribuídos e dados públicos, utilizando tecnologias da Web Semântica (WESTERSKI; IGLESIAS, 2011).

No artigo *Linked opinions: describing sentiments on the Structured web of data (2011)*, apresenta-se uma ontologia de opinião, demonstrando os benefícios da publicação na Web e os resultados da mineração de opinião de uma forma estruturada. Uma solução para descrever opiniões da web com padrões de metadados bem conhecidos e difundidos da Web Semântica. Ocorre ainda a demonstração da forma de adaptação da especificação de metadados disponível para auxiliar na vinculação de opiniões com outros conceitos sobre a web, levando a melhores recursos de pesquisa e uma melhor exposição dos dados (WESTERSKI; IGLESIAS; RICO, 2011).

No trabalho intitulado de *Application of semantic search in idea management Systems (2012)*, é apresentado um modelo, projeto e arquitetura baseado em busca semântica para sistemas de inovação aberta com foco em sistemas de Gestão de Ideias. Com a apresentação de uma metodologia para coleta, organização e busca de ideias, melhorando a interação entre usuários e simplificando o processo de análise de ideias (POVEDA; WESTERSKI; IGLESIAS, 2012).

Os quatro últimos trabalhos apresentados fazem parte de um mesmo projeto denominado GI2MO. A plataforma GI2MO é um sistema de gerenciamento de ideias de código aberto baseado na arquitetura de componentes do Sistema de Gerenciamento de Conteúdo Drupal® (POVEDA; WESTERSKI; IGLESIAS, 2012). O projeto GI2MO visa organizar todas as fases do processo de Gestão de Ideias. Ainda de acordo com o site do projeto, GI2MO possui como objetivo a configuração de tecnologias de Web Semântica no ambiente de Sistemas de Gestão de Ideias.

O projeto permite realizar busca automática, exploração do significado semântico para melhorar ideias e possibilita incorporar ideias através de *Linked Data*. O grande objetivo é a interoperabilidade com as soluções existentes. Oferece ainda formato semântico para as ideias de acordo com a GI2MO ontologia. A ontologia proposta pelo projeto GI2MO foi desenvolvida com o intuito de manter a integridade com as tendências e normas da Web Semântica e proporcionar uma ontologia simples, colocando impacto sobre sua usabilidade e facilidade (GI2MO, 2014).

De acordo com a homepage do projeto, o foco principal é desenvolver uma ontologia para modelar processos de Gestão de Ideias, visando alcançar resultados reais e mensuráveis. Atualmente, de acordo com a *homepage*¹ do projeto, entre as funcionalidades da plataforma estão um sistema de recomendação de ideias, métricas para avaliação da ideia, agrupamento de ideias, utilização de *linked data*, classificação de ideias, aplicação de busca semântica, análise de relacionamento e análise de sentimento. O projeto vem recebendo incentivos e evoluindo em publicações.

Ao analisar a ontologia proposta pelo projeto GI2MO, observou-se a existência de várias classes e subclasses descontinuadas, bem como, propriedades de dados e objetos. A ontologia no momento da análise não promovia suporte à formação de agrupamentos. Em face destas características optou-se por desenvolver uma ontologia com base nas características do modelo proposto por esta dissertação.

2.3 DESCOBERTA DE CONHECIMENTO

A descoberta de conhecimento representa o processo de identificar, receber informações relevantes e por meio desta, computá-las e agregá-las ao conhecimento prévio, alterando o estado do conhecimento atual, a fim de que determinada situação ou problema possa ser resolvido (WIVES, 2004).

O processo de Descoberta de Conhecimento em Bases de Dados (do inglês *Knowledge Discovery in Databases* – KDD) possui como objetivo identificar e desvendar relacionamentos implícitos entre informações armazenadas em bases de dados (SILVA; ROVER, 2011). Está área surgiu como uma alternativa para buscar soluções para o

¹ <http://www.gi2mo.org/>

problema de sobrecarga de dados na era da informação digital (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996).

Berry e Linoff (1997) definem o processo de KDD como a análise e a exploração automática ou semiautomática de grandes volumes de dados, cujo objetivo está em desvendar regras e padrões significativos. Após evidenciar padrões, estes são utilizados para auxiliar no processo de tomada de decisão no domínio de análise (CAO *et al.* 2010).

O objetivo principal deste processo está na descoberta de conhecimento interessante e potencialmente útil (VASHISHTHA; KUMAR; RATNOO, 2012). Para Fayyad, Piatetsky-Shapiro e Smyth (1996) este processo consiste na tradução de dados brutos em informações relevantes.

O processo de descoberta de conhecimento em bases de dados manipula informações estruturadas, em geral, armazenadas em bancos de dados relacionais e/ou orientado a objetos. Noack e Schmitt (2013) afirmam que o processo de KDD destina-se a facilitar e acelerar a extração de conhecimento a partir de fontes de dados persistentes.

No processo de KDD os dados estão estruturados e dispostos em formatos numéricos ou símbolos, informações estas compreendidas por sistemas computacionais (WEISS *et al.*, 2005).

Para Fayyad *et al.* (1997), o campo de pesquisa que envolve o KDD compreende o desenvolvimento de técnicas e métodos que buscam fornecer significado aos dados.

As etapas compreendidas no processo de KDD são: a seleção dos dados, o pré-processamento responsável por adequar os dados aos algoritmos, a mineração efetiva dos dados que abrange o uso de técnicas geralmente baseadas na Inteligência Artificial ou Estatística (MAIA; ROCHA, 2010), a validação dos resultados e a análise e interpretação dos resultados para aquisição do conhecimento.

Porém, a informação estruturada não compreende a estrutura adotada por todas as informações dispostas no contexto digital. Segundo Tan (1999), Feldman e Dagan (1995) e Kuechler (2007), cerca de 80% da informação produzida, seja em uma organização ou até mesmo por usuários comuns, encontram-se no formato textual decorrido da linguagem natural.

Em face desta afirmação, surgiu o processo de Descoberta de Conhecimento em Textos (do inglês *Knowledge Discovery in Texts* – KDT). O mesmo é similar ao processo de KDD, entretanto trabalha com uma coleção de documentos em linguagem natural, buscando padrões e

tendências, classificando e comparando documentos (SILVA; ROVER, 2011).

Conforme Tan (1999) e Feldman *et al.* (2001) tanto o KDT quanto o KDD referem-se ao processo de extração de padrões não triviais e de conhecimento útil. Entretanto, a área que envolve o KDT torna-se mais complexa devido à falta de estruturação da informação.

O autor Kuechler (2007) aponta que atualmente grande parte do conhecimento encontra-se no formato textual, e por este motivo esse conhecimento necessita ser identificado, representado e manipulado de modo a tornar-se potencialmente útil para as organizações. Na visão de Hearst (1999) os documentos textuais contêm informações preciosas, porém codificadas e difíceis de serem interpretadas automaticamente.

A informação textual possui uma estrutura que necessita da aplicação de técnicas especializadas para serem analisadas por meios computacionais, devido ao significado implícito atribuído a cada termo presente na linguagem humana (PALAZZO *et al.*, 2006; SABOL, 2009).

2.4 ANÁLISE DE AGRUPAMENTO

A literatura apresenta diferentes nomenclaturas para o termo Análise de Agrupamento, como Taxonomia Numérica, *Data Clustering*, Análise de grupos, Análise de Aglomerados (ou Análise de *Clusters*, do inglês, *Cluster Analysis*) e reconhecimento de padrões não supervisionados.

A análise de agrupamento corresponde a uma das operações mais utilizadas na exploração e análise de dados, focada na descoberta e interpretação de grupos de objetos tendo propriedades e/ou comportamentos semelhantes (ANDRIENKO; ANDRIENKO, 2009). Representa uma análise rotulada como aprendizado não supervisionado, não possuindo classificação prévia dos dados para posterior agrupamento (KONCHADY, 2006).

Para Cormack (1971), Hair *et al.* (2010) e Lattin, Carroll e Green (2003), a análise de agrupamento versa sobre a aproximação interna dos objetos (homogeneidade) e o distanciamento externo (separação) entre grupos. De modo geral, os algoritmos particionam um conjunto de objetos em aglomerações (MANNING; SCHÜTZE, 2003).

Segundo Jain, Murty e Flynn (1999) a análise de agrupamento consiste na organização de um conjunto de padrões (geralmente dispostos em vetores ou em espaços multidimensionais) em grupos, de

acordo com um grau de semelhança. Hair *et al.* (2010) também corroboram desta informação e salientam que o principal objetivo deste processo é situar as observações homogêneas em grupos, a fim de definir uma estrutura para os dados.

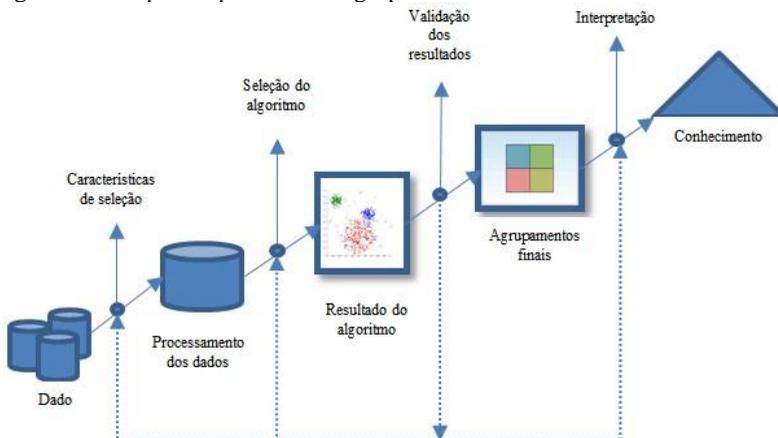
Por meio desta análise é possível determinar informações acerca do grau de semelhança ou diferença entre dois ou mais objetos (VASCONCELOS *et al.*, 2007), ou ainda, propiciar a descoberta não supervisionada de temas e dos principais tópicos a partir de uma coleção de documentos (ALJABER *et al.*, 2009).

Esta técnica é potencialmente útil, pois auxilia na análise de extensos volumes de dados, uma vez que permite a um analista considerar grupos de objetos em vez de objetos singulares. Para Andrienko e Andrienko (2009) agrupamentos também podem ser úteis para outros fins, como na detecção de objetos incomuns, que podem exigir uma investigação especial.

Sculley (2010) enfatiza a importância desta análise para agrupar resultados de busca, para a identificação de resultados duplicados e para agregação de conteúdos semelhantes. A visualização e interpretação dos agrupamentos por um especialista possibilita atribuir significado e valor a informação, podendo auxiliar no processo de tomada de decisão.

Existem diferentes métodos para realização do processo de agrupamento de dados. A distinção entre os métodos está relacionada ao tipo de resultado a ser obtido e pelas diferentes formas de definir a proximidade entre um indivíduo e um grupo já formado, ou entre dois grupos quaisquer (VASCONCELOS *et al.*, 2007).

Figura 11 - Etapas do processo de agrupamento



Fonte: Adaptado de Halkidi, Batistakis e Vazirgiannis (2001)

Para Halkidi, Batistakis e Vazirgiannis (2001) as etapas presentes no processo de agrupamento podem ser evidenciadas na Figura 11. Para os autores a etapa de Seleção das Características corresponde à seleção dos atributos para realizar o agrupamento. A Seleção do Algoritmo envolve a escolha do algoritmo propriamente dita, bem como a definição dos parâmetros, tais como a escolha da medida de proximidade. A etapa de Validação dos Resultados determina a qualidade do processo de agrupamento, e a etapa de Interpretação representa o momento no qual o especialista de domínio irá analisar os resultados e influir sobre padrões, evidências e tendências.

Considerando dados não estruturados, o processo de agrupamento inicia-se com o levantamento dos termos e/ou palavras-chave presentes em cada documento do conjunto de análise adicionando-se estes termos a um vetor. Cada vetor é responsável por identificar um documento. Posteriormente, determina-se a proximidade entre os documentos com base nos vetores criados na etapa anterior. A partir deste processo é gerado um valor que definirá se os documentos devem ou não ser agrupados (EBECKEN; LOPES; COSTA, 2003).

Os autores Han e Kamber (2001) afirmam que o processo de agrupamento necessita atender a determinadas características como:

- Escalabilidade;
- Manipulação de diferentes tipos de dados;
- Alta dimensionalidade;
- Descoberta de grupos com formas arbitrárias;
- Capacidade de tratar dados que possuam ruídos;
- Possuir conhecimento sobre o domínio para determinar os parâmetros de entrada;
- Não possuir sensibilidade em relação à ordem dos registros de entrada;
- Possibilitar a inclusão de restrições específicas do usuário;
- Permitir a interpretação e a usabilidade.

O resultado obtido com a aplicação do algoritmo de agrupamento pode ser influenciado pelo número de grupos a serem formados ou pela medida de similaridade ou dissimilaridade selecionada (GOWER; LEGENDRE, 1986; JACKSON; SOMERS; HARVERY, 1989).

2.4.1 Algoritmos hierárquicos

Na literatura as técnicas de agrupamento são classificadas em técnicas hierárquicas ou não hierárquicas. A técnica hierárquica envolve a construção de uma hierarquia baseada em árvore. Promove uma seqüência de partições, onde cada partição é agregada a partição vizinha na seqüência (JAIN; DUBES, 1988).

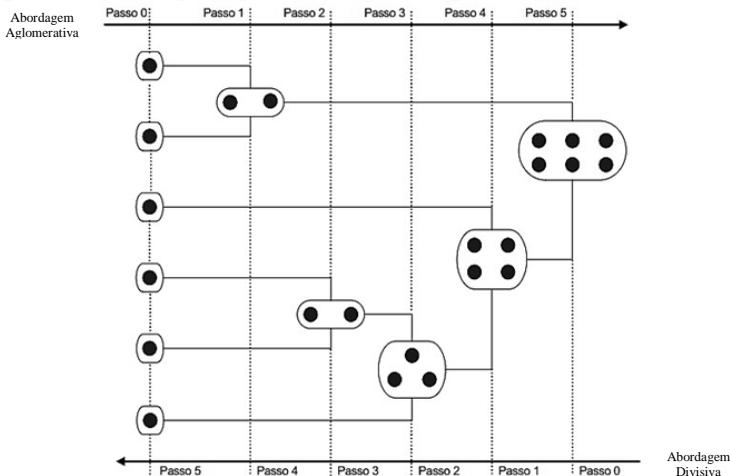
A abordagem hierárquica é considerada simples, promovendo divisões sucessivas entre os dados, originando uma representação no formato de hierarquia (EVERITT, 2001).

Esta técnica possui como vantagem o fato de não assumir um número predefinido de grupos. O resultado obtido com aplicação de técnicas hierárquicas pode ser apresentado por uma árvore de classificação denominada dendograma.

As técnicas hierárquicas são classificadas em duas abordagens: as aglomerativas e as divisivas. Estas abordagens apresentam como desvantagem o fato de ao atribuir um objeto a determinado grupo, este objeto não poder ser realocado para outro grupo (LATTIN; DOUGLAS; PAUL, 2011).

A técnica hierárquica aglomerativa ou divisiva fundamenta-se na edificação de uma matriz de similaridade (dissimilaridade) ou distâncias, onde cada elemento da matriz expõe o grau de diferença entre cada dois casos, de acordo com as variáveis escolhidas.

Figura 12 - Abordagem aglomerativa e abordagem divisiva



Fonte: Adaptado de Han e Kamber (2001)

A abordagem aglomerativa possui como objetivo reunir os objetos em grupos cada vez maiores, incluindo não somente os objetos, mas também os agrupamentos já formados (WIVES, 2004). Todavia, na abordagem divisiva os objetos são organizados em um único grupo que sofrerá sucessivas divisões, até cada objeto encontrar-se em um agrupamento separado (WIVES, 2004).

Nos algoritmos hierárquicos a forma de medir a similaridade pode ser classificada de três formas segundo Jain e Dubes (1988):

- Algoritmo hierárquico de ligação simples (do inglês *Single Linkage*);
- Algoritmo hierárquico de ligação média (do inglês *Average Linkage*);
- Algoritmo hierárquico de ligação completa (do inglês *Complete Linkage*).

No método envolvendo a ligação simples, também conhecido como método do vizinho mais próximo ou da distância mínima, o algoritmo monta os agrupamentos de acordo com a maior similaridade entre quaisquer objetos de dois grupos, unindo assim dois grupos similares, promovendo a homogeneidade (JAIN; DUBES, 1988).

Já no método de ligação média, a similaridade é obtida por meio da média de distância entre todos os objetos de dois grupos em questão (JAIN; DUBES, 1988). Enquanto que no método de ligação completa, conhecido como método do vizinho mais distante, a similaridade é obtida através da menor distância entre dois objetos de grupos distintos (JAIN; DUBES, 1988).

2.4.2 Algoritmos por particionamento

Os algoritmos por particionamento foram desenvolvidos para agrupar objetos em n grupos, definidos antecipadamente ou definidos durante a execução do processo (JOHNSON; WICHERN, 2007).

De acordo com Hair *et al.* (2010) este algoritmo designa objetos a agrupamentos, levando-se em conta a definição da quantidade de grupos. Segundo Fung (2001) os métodos por particionamento são extremamente mais rápidos que métodos hierárquicos.

Uma base de dados possui n objetos, onde k é o número de agrupamentos desejados, os algoritmos por particionamento irão gerar k partições empregando medidas de distância, objetivando manter alta similaridade entre os objetos de um mesmo agrupamento, enquanto que

os objetos deste agrupamento quando comparados com os objetos de outras partições apresentem baixa similaridade.

O primeiro passo é eleger uma semente como ponto inicial para formação de um agrupamento, e todos os objetos selecionados dentro de uma distância de referência pré-definida são inseridos no agrupamento resultante. Em seguida, outra semente de agrupamento é eleita, e o processo avança até realizar o agrupamento de todos os objetos.

Todavia, os algoritmos por particionamento possuem desvantagens, como o fato de elencar o número de agrupamentos a serem formados. Caso este número seja escolhido erroneamente, a cada iteração do algoritmo resultados diferentes podem surgir, o que poderá impor uma estrutura de dados, ao invés de identificar a estrutura inerente ao processo (FUNG, 2001; KAINULAINEN, 2002).

2.4.2.1 *K*-means

O algoritmo *k*-means é o mais conhecido algoritmo baseado em particionamento (JAIN; MURTY; FLYNN, 1999), consolidando-se como o mais utilizado no meio científico e industrial (BERKHIN, 2002). Este método foi introduzido por J. B. MacQueen em 1967. É um dos mais simples algoritmos de aprendizagem não supervisionada para lidar com o problema de agrupamento.

Primeiramente, o algoritmo inicia com a informação de quantos grupos serão formados durante o processo de agrupamento. Posteriormente, o algoritmo distribui um elemento para cada grupo. Inicialmente estes elementos serão a semente inicial de cada grupo e conseqüentemente o centroide. Durante a iteração do algoritmo, à medida que novos elementos forem atribuídos aos grupos, o centroide é recalculado, representando a média entre os elementos. O *k*-means utiliza geralmente a distância euclidiana para calcular a distância entre os elementos.

2.4.2.2 Lingo

O lingo é um algoritmo de agrupamento web, ou seja, baseado em resultados de pesquisa, que se preocupa em definir as melhores etiquetas (descrições) dos grupos. Segundo Osiński (2003) a ideia inicial do lingo é primeiramente encontrar descrições significativas de *clusters*, e em seguida, com base nas descrições determinar o conteúdo pertencente a

cada grupo. Os rótulos são identificados com base no espaço do modelo vetorial juntamente com a técnica de indexação semântica latente (do inglês, *Latent Semantic Indexing* – LSI), buscando identificar os conceitos presentes nos documentos.

Ainda de acordo com Osinski (2003) os passos do algoritmo consistem em primeiramente realizar o pré-processamento, onde para cada documento ocorrerá à filtragem de texto, a identificação do idioma do documento e a aplicação da técnica de *stemming* (identificação da raiz de uma palavra). Posteriormente é realizada a extração de características visando descobrir termos e frases frequentes. O terceiro passo visa à inferência do rótulo do *cluster* utilizando LSI para descobrir os conceitos abstratos e para cada conceito identificar a frase com melhor correspondência. Após este passo, retiram-se os rótulos de *clusters* semelhantes. O passo quatro objetiva a descoberta de conteúdo, onde para cada rótulo de *cluster* será utilizado o Modelo do Espaço Vetorial para determinar o conteúdo de fragmentação. Por fim, ocorrerá a formação final do *cluster* e a fusão de agrupamentos similares.

Segundo Osinski e Weiss (2004) a diferença do Lingo em relação a outros algoritmos está no momento da descrição do agrupamento, pois ele tenta descobrir o melhor termo que descreva determinado agrupamento. Dada a preocupação do algoritmo em definir a melhor descrição para o agrupamento e como o objetivo do trabalho é a promoção de uma melhor visualização do domínio de Gestão de Ideias, o Lingo apresenta-se como uma boa alternativa.

Seus principais pontos fracos são a incapacidade de realizar agrupamento hierárquico e tempo de processamento bastante lento, devido ao uso de decomposição singular do valor (do inglês, *Singular Value Decomposition* – SVD) (WROBLEWSKI, 2003).

2.4.3 Índice de modularidade

A modularidade é uma medida que objetiva estimar a qualidade do algoritmo de agrupamento, ou seja, a qualidade da divisão realizada na rede (NEWMAN, 2006), sendo proposta por Girvan e Newman (GIRVAN; NEWMAN, 2002). Para Blondel *et al.* (2008) por meio do índice de modularidade é possível definir as regras para divisão da rede em grupos.

A modularidade Q avalia as conexões das partições de um grafo segundo o estudo de sua configuração em grafos aleatórios. O índice

soma a probabilidade da existência das arestas entre pares de nós *intra-cluster*, subtraindo a soma das arestas *intra-cluster*, dado pela fórmula.

$$Q = \sum_i (e_{ij} - a_i^2)$$

onde e_{ij} representa a fração das arestas da rede atribuídas a um grupo i , e a_i^2 é esta mesma fração, porém considerando que as arestas são inseridas aleatoriamente.

Resultados próximos a 0 significam baixa probabilidade da rede estar decomposta em comunidades reais. Nesta linha, observa-se que quanto mais o resultado for positivo e distante de 0 (zero) (valores iguais ou superiores a 0.3), aumenta-se a chance de que tais agrupamentos não existam apenas ao acaso, ou seja, quanto maior for a modularidade de uma partição, maior a probabilidade de realizarem-se agrupamentos.

Se o grafo gerado possui apenas um *cluster* C, então $Q(C) = 0$, valores não correspondentes a 0 apontam desvios de aleatoriedade (NEWMAN, 2003b).

Caso o valor obtido com aplicação da equação $Q(C)$ for próxima a 1, o grafo sugere uma estrutura forte de formação de grupos. Porém, isto nem sempre ocorre. Segundo os autores Newman e Girvan (2004), em redes reais estes valores ocorrem na faixa entre 0,3 e 0,7; onde redes com alto valor de modularidade possuem conexões densas entre os nós dentro do agrupamento e conexões esparsas entre nós de diferentes agrupamentos.

2.4.3.1 Medidas de Centralidade

O problema de agrupamento de dados explorado em grafos vem sendo utilizado há décadas por meio de diferentes abordagens, devido à forma de visualização e a praticidade proporcionada pelas estruturas formadas pela rede. Ao analisar a estrutura do grafo, podem-se obter informações relevantes quanto às interações e os elementos que compõem a rede.

Uma rede é constituída por um conjunto de vértices unidos por arestas, formando pontos conectados. Os vértices podem possuir propriedades associadas, como por exemplo, valores numéricos (NEWMAN, 2003a). Segundo os autores Easley e Kleinberg (2010)

uma rede pode ser vista como uma coleção de objetos conectados entre si por *links*.

A caracterização das redes pode ser definida pelas medidas de centralidade, as quais identificam a disposição de um elemento em relação à sua posição na rede. Segundo Gama *et al.* (2012) a posição do elemento perante a rede não é fixa. Determinada hierarquicamente, a centralidade manifesta a ideia de poder. Portanto, quanto mais central a posição do elemento, maior será a sua iteração e a sua influência na rede (GAMA *et al.*, 2012).

A estatística apoia a análise de redes por meio de medidas de centralidade e densidade. As medidas de centralidade do indivíduo, foco desta dissertação, estão centradas na localização do vértice na estrutura do grafo. Esta medida corresponde aos nós mais importantes perante a rede, ou seja, a forma de quantificar esta importância.

Segundo Freeman (1979), o primeiro a tentar organizar as medidas existentes, as medidas de centralidade podem ser divididas em: grau, valência ou prestígio, proximidade e intermediação. Abaixo será realizada uma explanação sobre as três medidas básicas, sendo, a intermediação, a proximidade e o grau.

- **Intermediação:** a centralidade de intermediação (do inglês, *betweenness*) baseia-se em todos os caminhos mínimos entre os pares de vértices da rede, medindo assim a frequência com que determinado nó aparece no caminho mais curto entre dois nós em questão. A ideia central é de que um nó está presente em muitos caminhos entre outros dois nós. A centralidade de intermediação de um vértice v é definida como:

$$CB(v) = \frac{CB(v)}{[(n-1)(n-2)/2]}$$

onde $CB(v)$ indica a quantidade de menores caminhos que passam pelo vértice v ; $n-1$ denota o número de caminhos mínimos entre o nó x e o nó y que passam pelo vértice v e $n-2$ denota o número de caminhos mínimos entre o nó s e o nó t .

- **Proximidade:** a centralidade de proximidade (do inglês, *closeness*) é definida a partir da média das distâncias geodésicas (menor distância necessária para conectar dois nós quaisquer) do vértice v e todos os outros vértices alcançáveis por ele. Para os autores Cook, Emerson e Gillmore (1983) as medidas baseadas em proximidade

medem a proximidade de um ponto específico em relação aos outros pontos. Apresenta-se a seguinte equação:

$$C_c(i) = \left[\sum_{j=1}^N d(i, j) \right]^{-1}$$

onde i é o nó central, j corresponde a outro nó da rede, e $d(i, j)$ é a menor distância entre os dois nós. Nesta equação, as distâncias são invertidas após terem sido somadas. Vértices com maior centralidade de proximidade ocupam uma posição próxima do centro do grafo.

- Grau: a centralidade de grau é considerada a medida mais simples perante as demais. Possui como premissa básica: se um nó possui muitas conexões, então este nó é “o nó” ou “um dos nós” mais importante(s) na rede. Para Freeman (1979) é o número de vértices adjacentes ao nó e com o qual possui contato direto. Esta medida está baseada na ideia de que vértices que possuem grau maior podem estar em uma localização privilegiada na rede. A centralidade de grau de um vértice v é definida como:

$$CD(v) = \text{grau}(v)/(n-1)$$

onde, $\text{grau}(v)$ representa a quantidade de arestas conectadas a um nó específico e n o número de nós envolvidos nas ligações diretas. A centralidade de grau em redes direcionadas pode ser vista ainda de duas formas (*indegree*) e (*outdegree*). A forma *indegree* corresponde a todos os nós que incidem sobre o nó em questão e a forma *outdegree* corresponde a todos os nós que são incididos pelo nó em questão.

2.4.4 Trabalhos correlatos relacionados ao domínio de Gestão de Ideias e Análise de Agrupamento

Esta seção consiste na identificação de trabalhos relacionados ao domínio de Gestão de Ideias e a Análise de Agrupamento, a partir de um conjunto de artigos selecionados e analisados no ano de 2016 por meio das principais bases de dados: *Scopus*, *ACM*, *IEEE* e *Springer Link*.

Objetivando alcançar um panorama acerca da análise de agrupamento no cenário de Gestão de Ideias, a seguinte expressão foi utilizada como filtro de pesquisa: (“*idea management*” and “*cluster*”³⁸),

nos campos referentes a título, palavras-chave e resumo nas bases que dispunham deste recurso. Ao total foram identificados 3 artigos, entretanto somente 1 apresentava a análise de agrupamento como uma solução para o domínio de Gestão de Ideias.

No artigo intitulado de *Idea Relationship Analysis in Open Innovation Crowdsourcing Systems (2012)*, apresenta-se uma proposição de uma série de métodos para sumarização do conjunto de dados. O objetivo consiste em verificar a utilidade da aplicação de uma hierarquia de relacionamento para dados de inovação aberta armazenados em sistemas de Gestão de Ideias. Os autores propuseram uma nova hierarquia de relações e demonstraram que a sua utilização pode aumentar significativamente a quantidade de relações obtidas (WESTERSKI; IGLESIAS; GARCIA, 2012).

No trabalho de Paukkeri e Kotro (2009), intitulado de *Framework for Analyzing and Clustering Short Message Database of Ideas*, é apresentado o Note, um *framework* para analisar e agrupar mensagens curtas em um banco de dados de ideias. Note é um *Noteboard* eletrônico compartilhado, onde os funcionários de uma empresa podem escrever as suas observações, ideias e perguntas. Note utiliza o algoritmo *K-means* para realizar os agrupamentos com o objetivo de criar e atualizar a memória coletiva de uma organização. A ferramenta combina métodos estatísticos de mineração de texto com o diário de práticas de uma organização de uma nova maneira, mesclando diferentes fontes de informação. Na pesquisa realizada, Note foi a solução tecnológica que mais se aproximou do objetivo desta dissertação, apesar de utilizar diferentes fontes de informação e não apenas ideias. Porém, os resultados não foram evidenciados no artigo, assim como maiores detalhes sobre o *framework*.

Para tanto, esta dissertação apresenta um modelo baseado no algoritmo Lingo para processar ideias e realizar o processo de agrupamento. Segundo Osinski e Weiss (2004) a diferença do Lingo em relação a outros algoritmos está relacionada à descrição do rótulo do agrupamento, objetivando encontrar o melhor termo que identifique determinado agrupamento. Devido à preocupação com a promoção de uma melhor visualização do domínio de Gestão de Ideias, o Lingo apresenta-se como uma boa alternativa.

2.5 CONSIDERAÇÕES FINAIS

Este capítulo discutiu as técnicas relacionadas ao trabalho advindas das áreas de estatística multivariada e descoberta de conhecimento. Delineou-se com mais ênfase as técnicas e algoritmos de análise de agrupamento e a representação semântica por intermédio das ontologias, visto que essas temáticas promovem suporte ao modelo proposto (discutido no próximo capítulo), utilizado tanto na identificação de relacionamentos entre ideias, quanto na representação semântica deste conhecimento. Realizou-se ainda uma explanação sobre o domínio de Gestão de Ideias. Estes procedimentos objetivam fornecer insumos para aplicações de Engenharia e Gestão do Conhecimento, tais como a formação de grafos para a explicitação de conhecimento, e assim auxiliar ao especialista de domínio na tomada de decisão.

3 MODELO PROPOSTO

Neste capítulo será apresentado o modelo proposto. A apresentação refere-se ao modelo lógico, sendo que o mesmo detalhará a interação decorrente entre os módulos componentes da proposição.

3.1 DESCRIÇÃO DO MODELO

A competitividade presente no cenário de mercado atual vem promovendo a busca constante por ideias que fomentem a inovação. A Gestão de Ideias gera desafios quanto à avaliação dos dados, representando o ponto crítico do processo, devido ao volume de submissões ou ainda pela trivialidade das ideias. A adequada análise e a destinação de recursos para investimentos são questões a serem tratadas para o efetivo processo de gerir ideias, objetivando maximizar os recursos para investimentos.

Ideias que emergem da organização, seja no formato textual impresso e/ou digital, são caracterizadas como documentos textuais. Em decorrência deste fato, possui-se como fonte de informação primária um conjunto de textos.

O modelo proposto permite que um conjunto de ideias seja agrupado de acordo com as características e semelhanças. Considera-se a formação de grupos a partir de documentos textuais.

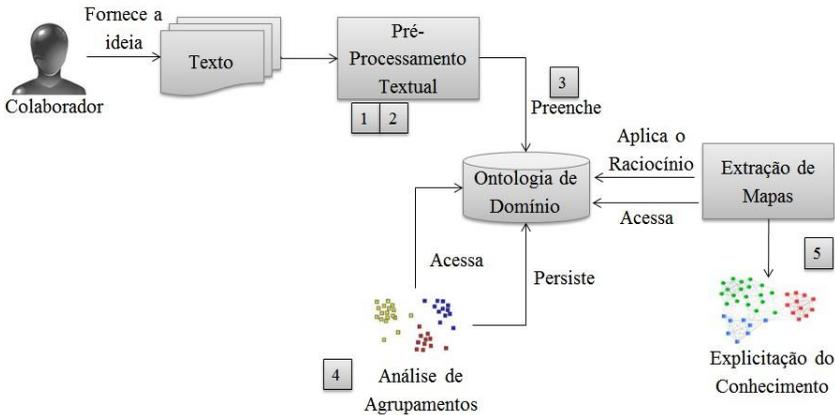
O foco deste trabalho é facilitar o processo de seleção e avaliação das ideias, permitindo a partir de um conjunto de ideias, identificar ideias adequadas, ou seja, passíveis de implementação e investimentos, auxiliando assim os gestores na tomada de decisão.

O modelo proposto está dividido em cinco etapas fundamentais que visam fornecer suporte ao processo de gerir ideias:

- 1ª etapa: a estruturação do conjunto de documentos;
- 2ª etapa: a indexação desta informação;
- 3ª etapa: o preenchimento da base de conhecimento;
- 4ª etapa: a realização do processo de agrupamento;
- 5ª etapa: a avaliação e explicitação do conhecimento.

Abaixo é apresentada a Figura 13 que representa o modelo proposto em que se pode observar como cada etapa está conectada.

Figura 13 - Modelo proposto



Fonte: Autora

As próximas subseções apresentam o detalhamento das etapas fundamentais deste modelo, sendo discutidas e justificadas as tecnologias utilizadas para a sua implementação.

3.1.1 Construção da estrutura de dados para processamento pelo modelo

A primeira etapa do modelo proposto corresponde à criação de uma estrutura de dados para ser processada pelo modelo proposto na dissertação.

O objetivo é realizar a estruturação dos campos presentes no documento textual fornecido pelo colaborador da organização, por meio da criação de documentos no formato XML. Contudo a descrição da ideia permanece sendo não estruturada. Esta fase possui como intuito adequar as ideias propostas em uma linguagem compreensível aos meios computacionais.

A Figura 14 apresenta o formato como as ideias estão dispostas no site da empresa Starbucks®, um dos cenários de estudo utilizados para avaliar a efetividade do modelo.

Figura 14 – Formato de exposição das ideias no site da empresa Starbucks®

Vote
Show Offers under Messages or Rewards - Not Account History
Share

Posted on 4/13/2015 8:57 AM
by werling

I can't tell you the numerous times I've miss out on using offers because I never saw them. How many people check their account history daily? I don't. So my offers go unnoticed until I check my history, usually monthly, and they have expired an I'm pissed. Why can't you put my offers under messages or even along with my rewards so I will notice and use them. I can't use something I don't know exists.

Comments [2]

Hide Comments [2]

Screen__Name

4/13/2015 11:39 AM

How many people check their history daily? Well, depending on which suggestions you've read in these discussions, it seems quite a few people review their history continually, and want even more detailed history posted.

h3t3pu 4us4r

5/13/2015 8:08 PM

Bonus - If the offers come as Messages, we will receive app notifications for them. Account History doesn't give app notifications.

Fonte: Site da Starbucks®²

Na figura encontra-se o título da ideia, a data e hora de postagem da ideia, o nome do autor, a descrição da ideia, *links* para compartilhamento da ideia. Do lado esquerdo está a quantidade de votos que a ideia recebeu. Votos podem ser positivos como no caso acima, ou negativos. Votos negativos são representados com o sinal de menos. Na figura têm-se ainda os comentários indicando o nome do autor do comentário e a descrição do mesmo, bem como a data e hora de postagem.

A Figura 15 apresenta o documento criado no formato XML para representar ideias com o objetivo de estruturar a informação. Pode-se observar que são extraídas algumas características das ideias, entre elas: a) o título da ideia; b) o dia, o mês e o ano em que a ideia foi sugerida; c) o nome do(s) autor(es); d) a categoria em que a ideia esta incluída; e) as palavras-chave; f) o número de votos; e g) o(os) comentário(s) que a

² <http://mystarbucksidea.force.com/apex/idealist>

ideia recebeu. Cada comentário poderá ter ainda um valor de avaliação (*rating*) e o nome do autor do comentário.

Figura 15 - Documento XML

```
<DOCUMENT
  ID = ""
  TITLE = ""
  DAY = ""
  MONTH = ""
  YEAR = ""
  IDEA = "" >

  <AUTHORS>
    <ITEM NAME = ""/>
  </AUTHORS>

  <CATEGORIES>
    <ITEM DESCRIPTION = ""/>
  </CATEGORIES>

  <KEYWORDS>
    <ITEM DESCRIPTION = ""/>
  </KEYWORDS>

  <VOTES AFFIRMATIVE = "" NEGATIVE = "" NEUTRAL = ""/>

  <COMMENTS>
    <ITEM COMMENT = "" NAME = "" RATING = ""/>
  </COMMENTS>
</DOCUMENT>
```

Fonte: Autora

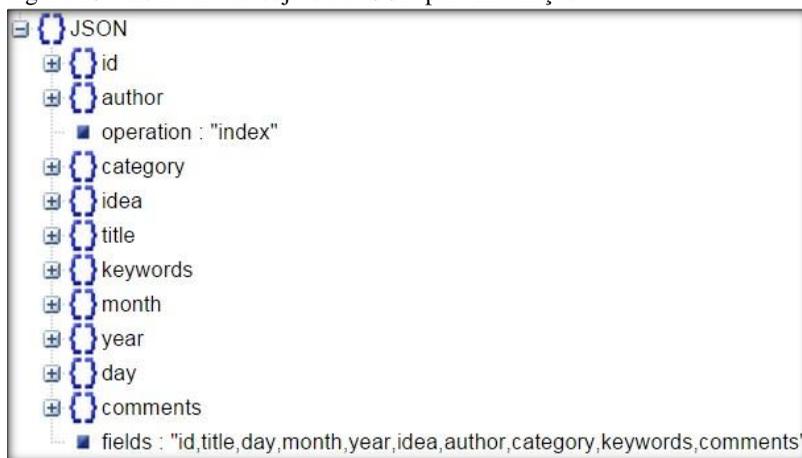
3.1.2 Indexação das ideias

Na segunda etapa, o sistema possui como entrada uma coleção de documentos no formato XML para representar as ideias fornecidas em geral por colaboradores da organização. Esses documentos são convertidos em um índice textual que facilita o acesso às informações contidas nos documentos da coleção.

O modelo baseia-se no *framework* (também referenciado como uma biblioteca) de indexação de documentos Lucene®, que possui características como escalabilidade e *software* livre, criado a partir da organização Apache Software Foundation®. O Lucene® é um *framework* de alto desempenho para busca textual escrito na linguagem de programação Java® e de fácil aplicabilidade.

A Figura 16 apresenta a estrutura do objeto JSON³ criado para ser repassado ao *framework* Lucene® para indexação das ideias. Segundo o site do projeto, um objeto JSON (*JavaScript Object Notation* - Notação de Objetos JavaScript) significa uma formatação para troca de dados, onde seres humanos interpretam e escrevem de forma facilitada, e máquinas compreendem e geram objetos mais facilmente. O JSON é um formato de texto completamente independente de linguagem, pois usa convenções que são semelhantes às linguagens C e familiares, incluindo C++, C#, Java, entre outras (JSON, 2016).

Figura 16 – Estrutura do objeto de JSON para indexação



Fonte: Autora

Como pode ser observada na imagem, a estrutura apresenta as mesmas características presentes no documento XML.

3.1.3 População da ontologia

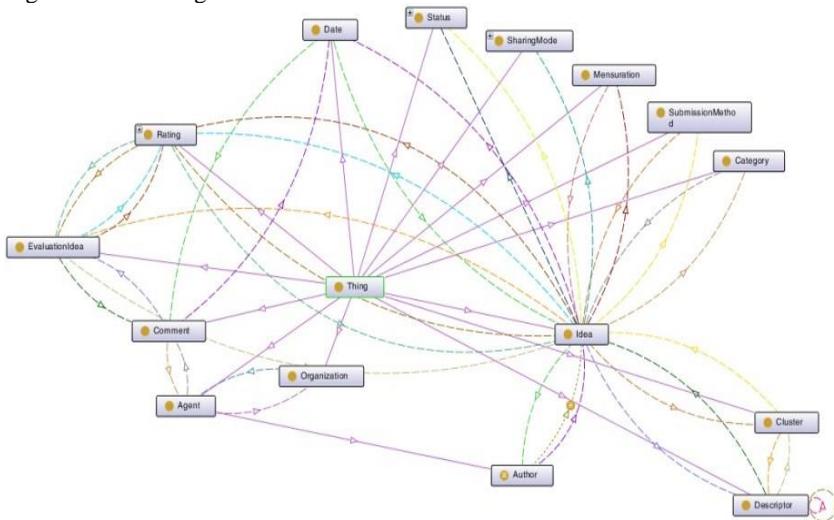
Em fase posterior ao processo de indexação, etapa 3 do modelo, ocorre a população da ontologia de domínio voltada à Gestão de Ideias. A ontologia foi desenvolvida como forma de representação do conhecimento de domínio. Não se identificou na literatura e em pesquisas realizadas na web, nenhuma ontologia com suporte ao processo de agrupamento. Apenas ontologias para representar ideias.

³ <http://www.json.org/json-pt.html>

A Figura 17 apresenta a ontologia de domínio desenvolvida. Entre as classes principais presentes na ontologia encontram-se a classe que representa a ideia em si (Classe *Idea*) e a classe *Cluster*, responsável por representar o resultado do processo de agrupamento. As principais propriedades de dados da classe *Idea* são: o título da ideia, a descrição da ideia, a data de submissão da ideia e um identificador.

A classe *Cluster* contém um identificador, a descrição do agrupamento, e através da propriedade de objetivo *hasIdea* são conectadas todas as ideias que estão presentes em determinado grupo. Esta conexão é realizada após a execução do processo de agrupamento de ideias.

Figura 17 - Ontologia de domínio



Fonte: Autora

A Figura 18 apresenta as classes presentes na ontologia de domínio. A classe *Agent* representa os colaboradores da organização que inseriram comentários nas ideias propostas. Possui ainda uma subclasse *Author*, responsável por representar os colaboradores que contribuíram com sugestões de ideias.

A classe *Category* é responsável por descrever as categorias que uma ideia está inserida. A classe *Comment* descreve os comentários que determinada ideia recebeu. A classe *Date* representa a data que determinada ideia foi inserida no sistema. A classe *Descriptor* descreve as palavras-chave das ideias.

A classe *EvaluationIdea* representa as avaliações que determinada ideia recebeu. Esta classe utiliza o conceito de reificação, ou seja, representando relações como objetos. O conceito de reificação foi necessário devido a uma única ideia possuir vários comentários. Via apenas uma propriedade de objeto não seria possível atribuir mais de um comentário e conseqüentemente as características do comentário à ideia. Através da utilização do conceito de reificação com a criação da classe *EvaluationIdea* foi possível criar a propriedade de objeto *hasComment*, responsável por conectar a classe *Idea* a classe *Comment*, quantos vezes for necessário, em decorrência do número de comentários que uma ideia recebeu.

A classe *Mensuration* representa o valor de investimento para concretização da ideia. A classe *Organization* é responsável por descrever as características da organização. A classe *Rating* representa o número de votos que a ideia recebeu. Os votos podem ser positivos *ratingUp* ou negativos *ratingDown*.

A classe *SharingMode* representa o modo de compartilhamento da ideia, que poderá ser público ou privado, por exemplo. A classe *Status* é responsável por descrever as fases que determinada ideia se encontra, como por exemplo, em fase de implementação, avaliação, entre outras.

Por fim, a classe *SubmissionMethod* representa o modo de submissão da ideia. Ideias podem ser submetidas, por exemplo, online, ou ainda por meio das tradicionais caixas de sugestões.

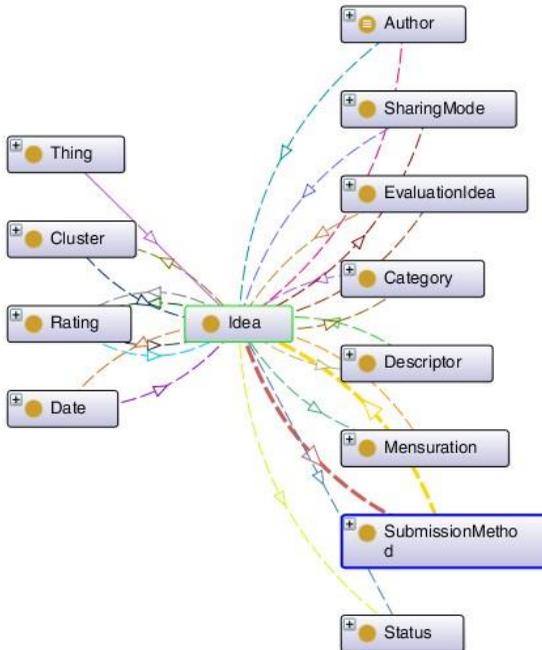
Figura 18 – Classes presentes na ontologia de domínio



Fonte: Autora

A Figura 19 apresenta a Classe *Idea*, uma das classes principais da ontologia e o seu relacionamento com as demais classes da ontologia. Esta classe se conecta via propriedades de objeto as classes *Author*, *SharingMode*, *EvaluationIdea*, *Category*, *Descriptor*, *Mensuration*, *SubmissionMethod*, *Status*, *Rating*, *Cluster* e *Date*.

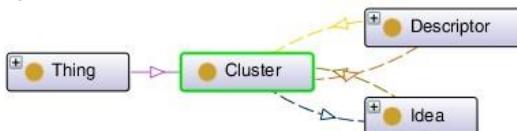
Figura 19 – Classe *Idea* e seus relacionamentos



Fonte: Autora

A classe *Cluster* também é uma das classes principais da ontologia, ao qual representa as informações obtidas por meio do processo de agrupamento. Esta classe possui relações com as classes *Descriptor* e *Idea*.

Figura 20 – Classe *Cluster* e seus relacionamentos



Fonte: Autora

3.1.4 Aplicação do algoritmo de agrupamento

Na fase de agrupamento, etapa 4, foi utilizado o algoritmo Lingo disponível no projeto Carrot²[⁴]. Segundo o site do projeto, o Carrot² é uma biblioteca com um conjunto de funcionalidades para construir um mecanismo de agrupamento de resultados de pesquisa, cujo objetivo é organizar os resultados da pesquisa em tópicos, de forma totalmente automática e sem conhecimentos externos, como taxonomias ou conteúdos pré-classificados.

Para os autores Osinski e Weiss (2004) a diferença do Lingo em relação a outros algoritmos está no momento da descrição do agrupamento, pois ele tenta descobrir o melhor termo que descreva determinado agrupamento. Dada a preocupação do algoritmo em definir a melhor descrição para o agrupamento e como o objetivo do trabalho é a promoção de uma melhor visualização do domínio de Gestão de Ideias, o Lingo apresenta-se como uma boa alternativa.

Basicamente a heurística do algoritmo Lingo é composta pela inversão do modo de agrupamento tradicional de documentos. Primeiramente, o algoritmo identifica os rótulos de fragmentação e numa etapa subsequente atribui os documentos aos rótulos para formar os grupos. Para determinar os rótulos, o algoritmo Lingo constrói uma matriz termo-documento para todos os documentos de entrada e decompõe a matriz para obter um número de vetores de base que mais se aproximam da matriz em um espaço de baixa dimensão. Cada um destes vetores dá origem a uma etiqueta de agrupamento. Para completar o processo, a cada rótulo é atribuído os documentos que contêm as palavras presentes no rótulo.

Para aplicação do algoritmo no modelo proposto, foi utilizada a versão disponibilizada pela biblioteca do projeto Carrot², desenvolvida em Java®. Como entrada do processo de agrupamento é necessário informar um termo, ou seja, uma semente para consulta (*query*) inicial.

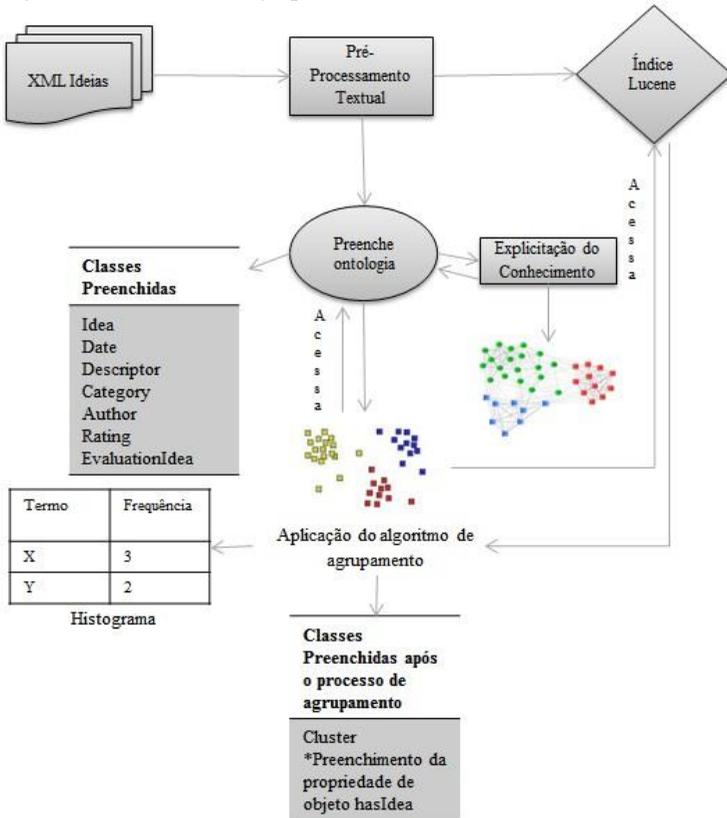
Para a identificação do termo inicial (semente) a abordagem utilizada no modelo proposto foi a criação de um histograma. O histograma é composto pelos termos mais frequentes presentes no conjunto de análise (ideias) desconsiderando *stop words* (termo com pouca capacidade discriminatória, tais como, artigos e preposições) e a frequência com que cada termo aparece neste conjunto. Ou seja, as sementes de pesquisa serão os termos que mais aparecem no conjunto em análise. As sementes são fundamentais para o processo, pois servirão

⁴ <http://project.carrot2.org/>

de rótulos iniciais para a formação dos agrupamentos. Após a formação do histograma, o conjunto é ordenado de maneira decrescente pelo valor de frequência, ou seja, os termos mais citados ficaram no topo. A partir disto, são considerados os termos mais frequentes (mais citados) respeitando um determinado limiar que então servirá de entrada para o algoritmo.

A Figura 21 ilustra os passos fundamentais descritos para realização do processo de agrupamento baseado no algoritmo Lingo:

Figura 21 - Processo de agrupamento



Fonte: Autora

Para iniciar o processo de agrupamento, o protótipo acessa o índice gerado, recupera as informações e inicia o processo iterativo de agrupamento com base nas sementes de consulta “*query*”. Todos os

A rede é formada por arestas que representam as ligações entre os nodos, e os próprios nodos que representam os agrupamentos através dos rótulos mais representativos de um conjunto de ideias. No cenário expresso na Figura 22 foram identificados quarenta documentos contendo o termo de consulta “*App*”, sendo divididos em dez grupos. Relacionado ao domínio de “*App*” pode-se perceber que os colaboradores da organização requerem aplicativos relacionados ao domínio de “*Drinks back in the new app*”, representado pelo grupo na cor marrom, sugerindo que os *drinks* favoritos dos usuários sejam salvos no *app*; o desenvolvimento de aplicativos para “*Windows*”, apresentado pelo grupo na cor azul escuro e aplicativos para usuários de “*Windows phone*”, representado na cor roxo, sendo este o grupo mais representativo no exemplo.

Ainda no grupo “*Drinks back in the new app*”, ocorre um forte apelo por aplicativos relacionados à questão nutricional, como informações nutricionais para aplicativos para Iphone® e a questão do cálculo nutricional.

3.2 CONSIDERAÇÕES FINAIS

Este capítulo apresentou o modelo proposto detalhando cada uma das etapas que o constitui. Os pontos centrais do modelo são compostos pela realização do processo de agrupamento e pela representação semântica das ideias. Adicionalmente, o modelo incluiu outras etapas para que os objetivos do trabalho sejam alcançados, entre elas, preparação dos dados, a indexação das ideias e a explicitação do conhecimento de domínio. Nesse sentido, a integração dessas técnicas em um único modelo visa fornecer um ferramental que promova suporte às aplicações de Engenharia e Gestão do Conhecimento voltado a Gestão de Ideias.

4 APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS

Neste capítulo serão apresentados e discutidos os principais resultados obtidos com o modelo. A principal motivação para implementação do modelo voltado a Gestão de Ideias, foi assegurar que ele funcionasse e alcançasse os objetivos. Na implementação considerou-se um conjunto de ideias coletadas do site de duas grandes empresas, no caso Starbucks® e Dell®. Dessa forma, três cenários foram elaborados com o objetivo de demonstrar a efetividade do modelo para auxiliar na tomada de decisão no domínio de Gestão de Ideias.

Os passos que envolvem a implementação do modelo proposto através de um protótipo são: pré-processamento dos documentos, indexação dos documentos, a construção dos histogramas, a criação da base de conhecimento, a aplicação do processo de agrupamento e a explicitação e avaliação do conhecimento.

4.1 CENÁRIOS DE ESTUDO

O primeiro cenário construído envolveu a coleta aleatória de 55 (cinquenta e cinco) ideias no domínio de novas tecnologias, presentes no site da empresa Starbucks®⁵. O objetivo é evidenciar relações e tendências abrangendo o cenário de mercado da empresa. A escolha da base se deu por motivos de estruturação do formato de disposição das ideias, bem como a mesma ter sido utilizada em outros cenários de estudo envolvendo o domínio de Gestão de Ideias, como no trabalho de Westerski, Dalamagas e Iglesias (2013). O período de coleta e extração dos dados para compor a base de dados ocorreu no final do ano de 2014.

A empresa Starbucks®, multinacional fundada em 1971, possui a maior cadeia de cafeterias do mundo com sede em Seattle, EUA. Atualmente, com mais de 15 mil lojas em 50 países concretiza-se como uma empresa de nível mundial de torrefação e venda de café diferenciado.

O segundo cenário objetivou aplicar o modelo proposto em um conjunto maior de ideias, com o intuito de avaliar a aplicabilidade e robustez do mesmo. Para tanto, foram coletadas 413 (quatrocentos e treze ideias), compreendidas entre o período de outubro de 2014 a março de 2015, do site da empresa Starbucks®. O domínio também correspondeu ao de novas tecnologias. Neste domínio existem mais de

⁵ <http://mystarbucksidea.force.com/apex/idealist?lsi=0>

5600 (cinco mil e seiscentas) ideias disponíveis. Dentre os domínios disponíveis, o que possui mais ideias é o de *Coffee & Espresso Drinks*, que representa o foco da empresa, com mais de 45000 (quarenta e cinco mil) ideias.

O terceiro cenário foi construído com base no site da empresa Dell®. A Dell® é hoje umas das maiores fabricantes de *hardware* do mundo, sendo sua sede nos Estados Unidos. No site da empresa foram coletadas 100 ideias, compreendidas entre o período de outubro de 2014 à março de 2015, correspondente ao período de 6 meses. O domínio escolhido foi “todas as categorias”.

4.1.1 Preparação dos dados

Primeiramente, visando o processo de estruturação das ideias, foram criados documentos no formato XML para cada uma das ideias coletadas. A Figura 23 apresenta a estruturação de uma ideia coletada para o primeiro cenário de estudo do site da empresa Starbucks®.

Figura 23 – Exemplo de uma ideia estruturada no formato XML

```
<DOCUMENT
  ID = "46"
  TITLE = "App updates"
  DAY = "19"
  MONTH = "07"
  YEAR = "2014"
  IDEA = "I was disappointed when I discovered that after updating my app several features went away,
  like being able to store mine and friend favorite drinks and being able to personalize a drink and getting the
  nutritional information after doing so. Usually an update means but fixes, improvements to existing features and
  new features or functionality. It is uncommon to lose features and functionality and it left me wishing I hadn't updated the app.">

  <AUTHORS>
    <ITEM NAME = "3McsMom"/>
  </AUTHORS>

  <CATEGORIES>
    <ITEM DESCRIPTION = "New technology"/>
  </CATEGORIES>

  <KEYWORDS>
    <ITEM DESCRIPTION = "App"/>
    <ITEM DESCRIPTION = "Updates"/>
  </KEYWORDS>

  <VOTES AFFIRMATIVE = "40" NEGATIVE = "" NEUTRAL = "">

  <COMMENTS>
    <ITEM COMMENT = "Several features didn't go away when you updated your app. The new app simply was simplified
    and several features were deemed to not have enough value over simplifying the app. You wrote It is uncommon to lose features
    and functionality. Have you thought that what Starbucks did was intentional and that it didn't want to have a bloated app that
    simply keeps adding things to the app, rather than making the app as simple as can be, which apparently is what a majority of users
    wanted?" NAME = "Screen__Name" RATING = ""/>
  </COMMENTS>
</DOCUMENT>
```

Fonte: Autora

Após a anotação manual dos documentos, passou-se para a segunda etapa do pré-processamento, ou seja, a extração via aplicação dos dados contidos nos documentos XML e a criação de um objeto

JSON contendo estes mesmos dados. Após a criação do objeto JSON, o mesmo é repassado a um servidor de indexação, no caso Lucene Apache®, que com base no objeto irá realizar a indexação dos documentos.

Com a indexação dos documentos contendo as ideias finalizada, a aplicação realiza o processo de inserção dos dados na base de dados, base esta correspondente a ontologia de domínio voltada a Gestão de Ideias.

4.1.2 Algoritmo de agrupamento

A partir da análise de agrupamento é possível explorar o conteúdo de extensos volumes de dados, possibilitando ao especialista de domínio analisar grupos de objetos e não somente um único objeto. A análise de agrupamento auxilia ainda na detecção de objetos incomuns e os resultados podem ser utilizados para tomada de decisão empresarial.

O modelo proposto no trabalho promove suporte a identificação das principais e mais coocorrentes ideias presentes no domínio de análise, facilitando assim a visualização da base de conhecimento.

Com o término do processo de indexação e população da ontologia de domínio, inicia-se o processo de agrupamento dos dados. O processo de agrupamento, baseado no algoritmo Lingo, acessa o índice criado, gera um histograma com os termos mais frequentes no domínio em análise e após realiza a formação dos agrupamentos.

Todos os dados obtidos com a aplicação do algoritmo: identificação do agrupamento, nome do agrupamento e o peso do agrupamento são armazenados na ontologia de domínio. Ocorre também a conexão por meio da propriedade de objeto *hasIdea*, entre o agrupamento gerado e a ideia contida no agrupamento, significando que determinada ideia está associada/pertence ao referido agrupamento. Com estes dados armazenados na ontologia forma-se a base de conhecimento, permitindo que um especialista de domínio realize posterior avaliação.

A pesquisa visa demonstrar a efetividade do modelo na formação de agrupamentos por meio do algoritmo Lingo. O algoritmo Lingo mostrou-se adequado devido à sua capacidade de descrever os rótulos dos agrupamentos.

Ao analisar as informações expressas na formação dos agrupamentos, ocorre uma melhora no entendimento do domínio de

identificação de oportunidades e na tomada de decisão no momento da destinação de recursos para investimentos em ideias.

Para o início do processo de agrupamento é necessário fornecer as sementes que permitem a realização das consultas na base de documentos indexados. Por meio dessas sementes o Lingo gera os agrupamentos necessários. As sementes foram obtidas através da formação de um histograma baseado nos termos mais significativos do conjunto de análise. Para compor o histograma foram removidas as pontuações e as *stopwords*, bem como, tornar todas as palavras minúsculas e no singular. A Tabela 1 apresenta os cinco termos mais significativos do conjunto.

Tabela 1 - Termos mais frequentes no primeiro cenário de estudo

<i>Termo</i>	<i>Número de vezes em que aparecem nos documentos (ideias)</i>
<i>App</i>	38
<i>Phone</i>	20
<i>Drink</i>	18
<i>Store</i>	15
<i>Windows</i>	15

Fonte: Autora

O termo mais significativo do conjunto é “*App*”, porém apesar dos termos “*Phone*” e “*Windows*” aparecerem separados na tabela, os dois representam o segundo par de termos que mais ocorre na análise.

Por intermédio das Redes Sociais Virtuais, caracterizadas como grupos ou espaços específicos na Internet, dados e informações podem ser compartilhados entre pessoas, sendo estes dados ou informações de caráter geral ou específico, disponibilizados nas mais diferentes formas (fóruns, comunidades, entre outros). Pode ocorrer ainda, a formação de grupos por afinidade gerando comunidades virtuais, seja em espaços abertos ou não, contribuindo em discussões, debates e apresentação de temas variados. Nesta linha e considerando este conceito, grandes empresas como Canonical®, Starbucks®, Dell®, tem incentivado e estimulado a criação de uma página no próprio site da empresa para que ocorra a formação de uma rede de colaboradores para fornecer ideias, ou comentários para ideias já postadas no site.

Com isso, e com a aplicação dos conceitos presentes na análise de redes sociais, crescente nas últimas décadas, é possível visualizar a formação das redes de forma clara e visual. De modo a atingir este

objetivo foi utilizado para visualização da formação dos agrupamentos a API Gephi®.

Por meio da API Gephi® é possível analisar claramente as sub-redes dentro da rede composta por ideias, permitindo explorar o fluxo das informações e os relacionamentos existentes entre as mesmas.

Segundo o site do projeto⁶, o Gephi® é uma ferramenta para exploração e compreensão de grafos voltada à análise exploratória de dados, onde o usuário interage com a representação e manipulação das estruturas, formas e cores, objetivando revelar propriedades ocultas do conjunto de análise (GEPHI, 2015).

Segundo Gephi (2015) “o objetivo é ajudar os analistas de dados a fazerem hipóteses, intuitivamente descobrir padrões, isolar singularidades de estruturas ou falhas durante a terceirização de dados”. O Gephi permite utilizar algoritmos baseados em força ou algoritmos baseados em multi-nível para compor a distribuição entre os nós. O mesmo permite utilizar métricas para análise de redes sociais e redes de escala livre, entre outras características, tais como:

- Intermediação, Proximidade, Diâmetro, Coeficiente de *Clustering*, caminho mais curto médio, *PageRank*, HITS;
- Detecção de Comunidades através do índice de modularidade;
- Geradores aleatórios de dados.

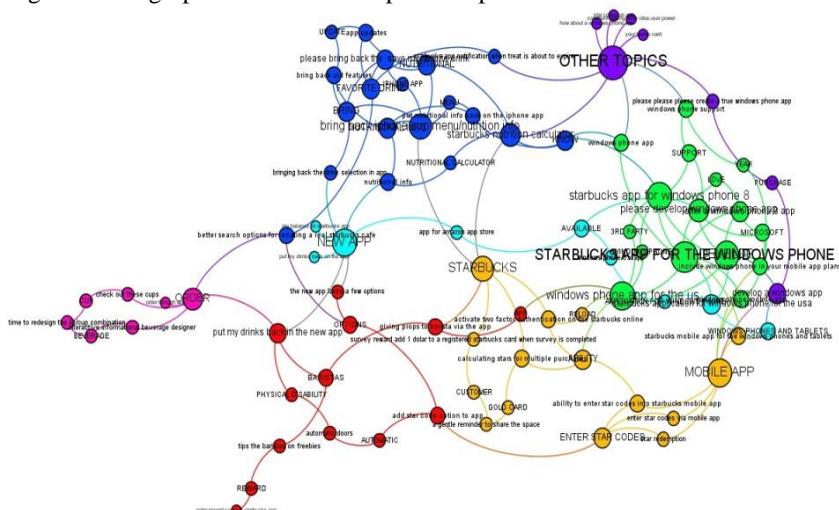
Para compor a visualização da rede foram utilizados como parâmetros de entrada no Gephi®:

- Visão geral da rede: os parâmetros analisados foram o grau médio, o diâmetro da rede e a modularidade;
- Algoritmo de distribuição: utilizou-se o force atlas;
- Partição dos nós: realizada através do grau;
- Classificação dos nós: realizada através do índice de modularidade;

A Figura 25 apresenta os agrupamentos obtidos para o primeiro cenário de estudo com a aplicação do protótipo desenvolvido. Nesta figura é apresentado o agrupamento total do modelo. Os nodos mais representativos serão explanados posteriormente.

⁶ <http://gephi.github.io/>

Figura 25 – Agrupamento formado a partir do primeiro cenário de estudo



Fonte: Sérgio, Gonçalves e Souza (2015)

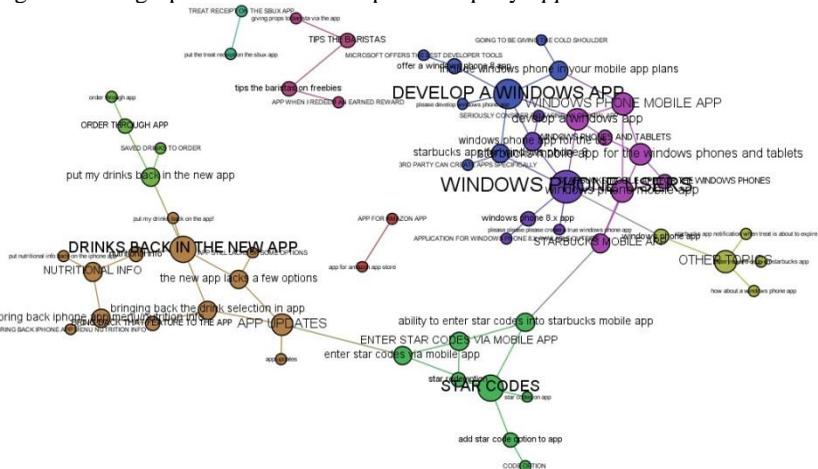
Após a formação da rede alguns dados podem ser obtidos, tais como o grau de modularidade, grau médio, diâmetro da rede, comprimento médio do caminho, entre outros. O valor da modularidade da rede nas condições apresentadas acima alcançou um índice de 0,636. Redes com o valor de modularidade alto possuem conexões densas entre os nós dentro do agrupamento e conexões esparsas entre nós de diferentes agrupamentos. Valores próximos a 1 indicam a forte presença de estrutura na rede. Em geral, os valores para a modularidade variam entre 0,3 e 0,7 (NEWMAN; GIRVAN, 2004). Em alguns casos podem chegar mais próximo de 1, mas são raros os casos em que isso acontece. Para o Grau Médio, que define o peso dos nodos de acordo com a quantidade de suas conexões, foi obtido valor igual a 2,696; o Diâmetro da rede, que significa a maior distância geodésica entre todos os pares de nodos presentes no grafo, atingiu valor 11; e o Comprimento Médio do Caminho foi de 5,049.

Pode-se observar que o nodo com maior representatividade é o de cor verde, denominado de “*Starbucks app for windows phone*”, seguido pelos nodos “*Other topics*” (cor roxo), “*Favorite drink*” (cor azul), “*Mobile app*” (cor amarela) e “*New app*” (cor azul). As palavras em maiúsculas correspondem à descrição do nodo formado a partir do algoritmo Lingo, e as palavras minúsculas ao título da ideia.

Ao analisar a formação dos agrupamentos, observa-se que a comunidade de colaboradores da empresa deseja aplicativos *mobile* para “*Windows phone*” principalmente na versão 8, evidência esta comprovada pelos nodos na cor roxo, azul escuro, verde, amarelo e azul claro. Pode-se ainda verificar que os colaboradores desejam aplicativos voltados para o termo “*Drinks*”. A questão nutricional também é um fator de destaque presente no nodo azul escuro, através do interesse por aplicativos que forneçam informações nutricionais ou ainda calculadoras nutricionais.

A Figura 26 apresenta o agrupamento obtido a partir da consulta à base de conhecimento em função do termo “*App*”.

Figura 26 – Agrupamento formado a partir da query *App*



Fonte: Sérgio, Gonçalves e Souza (2015)

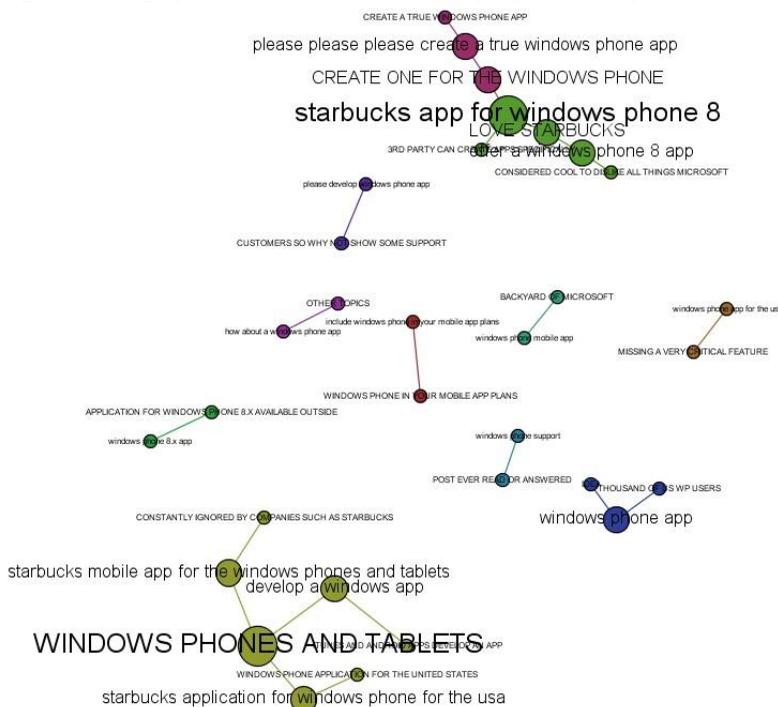
Na formação do agrupamento foram identificados quarenta documentos contendo o termo “*App*”, sendo agrupados em dez grupos. Relacionado ao contexto de “*App*” os colaboradores da organização solicitam mais aplicativos relacionados ao nodo identificado de “*Drinks back in the new app*”, representado na cor marrom, o desenvolvimento de aplicativos para “*Windows*”, apresentado pelo nodo na cor azul escuro e aplicativos para usuário de “*Windows phone*”, representado pelo nodo na cor roxo, sendo este o nodo mais representativo no domínio de análise.

Ainda referente ao nodo de “*Drinks back in the new app*”, o colaborador da ideia menciona que ele gostaria de salvar as suas

preferências no *app* e pedir os seus drinks favoritos já salvos previamente no aplicativo, devido à dificuldade na fala. Observa-se também neste grupo o forte apelo por aplicativos relacionados com a questão nutricional. Esta informação pode ser confirmada no que foi apresentado na Figura 25 contendo o agrupamento do conjunto de análise como um todo. As descrições do agrupamento são relacionadas a informações nutricionais para aplicativos para “*Iphone*” e a questão do cálculo nutricional.

Na Figura 27 apresenta-se o agrupamento obtido a partir da consulta à base de conhecimento em função do termo “*Windows phone*”.

Figura 27 – Agrupamento formado a partir da query *Windows phone*



Fonte: Sérgio, Gonçalves e Souza (2015)

Neste segundo cenário foram identificados quinze documentos contendo o termo “*Windows phone*”, sendo agrupados em onze grupos. Ocorre uma forte relação deste contexto com o contexto de aplicativos móveis, representado na Figura 27 pelo nodo na cor azul escuro com o termo “*Windows phone app*”.

Neste caso, observa-se que o nodo mais expressivo é representado pela cor verde, cuja descrição solicita o desenvolvimento de aplicativos para celulares que contenham o sistema operacional *Windows Phone*® da empresa Microsoft®, e o desenvolvimento de aplicativos para “*Tablets*”. Outra característica da formação deste agrupamento é que os colaboradores da empresa Starbucks® solicitam aplicativos para *Windows Phone*® para usuários dos EUA.

4.1.3.2 Segundo cenário de estudo

O segundo cenário de estudo envolveu também a empresa Starbucks® e foi baseado na coleta de 413 ideias. A Tabela 2 apresenta o histograma com os 8 termos mais frequentes no segundo cenário de estudo.

Tabela 2 - Termos mais frequentes no segundo cenário de estudo

<i>Termo</i>	<i>Número de vezes em que aparecem nos documentos (ideias)</i>
<i>App</i>	184
<i>Time</i>	107
<i>Store</i>	89
<i>Customer</i>	88
<i>Phone</i>	76
<i>Coffee</i>	72
<i>Card</i>	70
<i>Drink</i>	69

Fonte: Autora

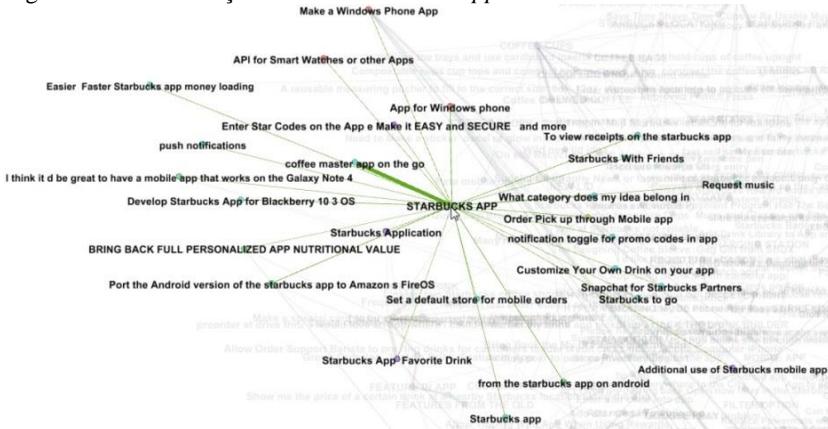
Na Tabela 2 pode-se perceber que independente do cenário anterior, novamente os termos “*App*”, “*Phone*” e “*Drink*” são altamente citados na plataforma da organização.

No momento da criação das redes contendo os agrupamentos, adicionou-se ao peso da aresta o valor que o Algoritmo Lingo gera no momento da execução do processo de agrupamento. Ou seja, cada nodo formado possui um peso específico perante o agrupamento como um todo. Quanto maior o peso do nodo, maior a significância do mesmo perante o conjunto. A Figura 28 apresenta a nuvem de termos com os termos mais citados no domínio de análise.

Foram gerados pelo processo de agrupamento um total de 260 nodos e 400 arestas. O valor da modularidade da rede nas condições apresentadas acima alcançou um índice de 0,758. Para o Grau médio, que define o peso dos nodos de acordo com a quantidade de suas conexões, foi obtido valor igual a 3,046; o Diâmetro da rede, ou seja, a maior distância geodésica entre todos os pares de nodos presentes no grafo, atingiu valor 12; e o comprimento médio do caminho foi de 4,961.

Em virtude do volume de dados optou-se por analisar os nodos com o maior valor correspondente ao item de avaliação “grau”. Para tanto, a Figura 30 apresenta o nodo com o maior valor correspondente ao grau, sendo este de 27, e os seus relacionamentos.

Figura 30 – Visualização do nodo *Starbucks app*



Fonte: Autora

No nodo denominado “*Starbucks app*” em virtude da espessura da aresta pode-se observar que os colaboradores da empresa desejam aplicativos voltados ao “*Coffee master app on the go*”, ou seja, um aplicativo relacionado a sugestão de um catálogo com os produtos que a empresa oferece, pois o usuário poderia ter acesso rapidamente a este conteúdo de forma mais prática.

A espessura da aresta, como mencionado anteriormente, é responsável por identificar o peso de determinado nodo no agrupamento. Quanto maior a espessura da aresta maior a representatividade do nodo no agrupamento. Novamente ocorre um apelo por aplicativos *mobile* voltados ao sistema operacional “*Windows phone®*” e “*Android®*”,

aplicativos voltados também ao celular da “*Blackberry 10.3 OS*” e “*Galaxy note 4*”, e aplicativos direcionados ainda a “*Favorite drinks*”. Algumas ideias interessantes como “definir uma loja padrão para encomendas móveis”, “um *app* personalizado com o valor nutricional”, “aplicativos para relógios inteligentes”, “notificações de códigos promocionais para os aplicativos”, “personalização da própria bebida via aplicativo”, “solicitar a ordem através do *app*”, “poder solicitar músicas”, “visualizar receitas no *app* da Starbucks”, estão presentes como sugestões para empresa na formação deste nodo. Enfim, observa-se que ocorre um apelo por aplicativos móveis.

A Figura 31 apresenta o nodo denominado de “*Windows phone*”. Este nodo possui grau no valor de 21 e, considerando as suas relações, representa o segundo nodo mais representativo do domínio de análise envolvendo o segundo cenário de estudo.

Figura 31 – Visualização do nodo *Windows phone*

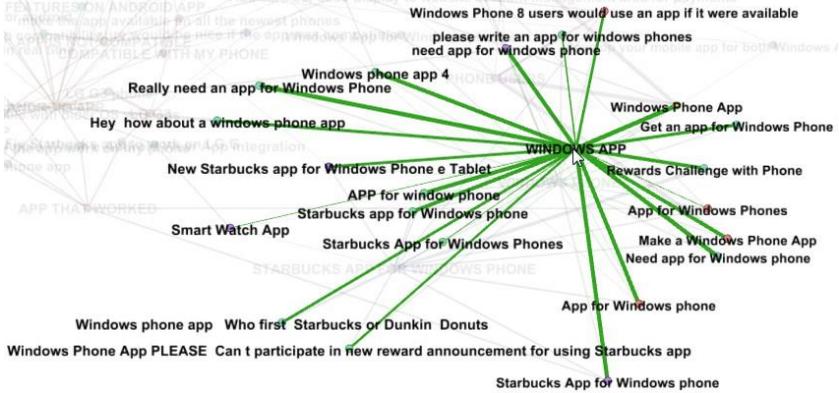


Fonte: Autora

Na figura acima se observa que relacionado ao contexto *Windows phone*® a um desejo por parte dos colaboradores por novos aplicativos para o sistema operacional da Microsoft® o *Windows phone*®, e também certa insatisfação por parte dos colaboradores. Uma das ideias interessantes presentes neste nodo é a possibilidade de adicionar a placa de vídeo para leitura do código de barras na área de gerenciamento de conta para pagamentos via *website*.

A Figura 32 corresponde ao nodo denominado de “*Windows app*”. O valor correspondente ao grau é de 21, compondo o terceiro nodo mais representativo do domínio de análise. A Figura 32 apresenta o nodo e seus relacionamentos.

Figura 32 – Visualização do nodo *Windows app*



Fonte: Autora

Assim como apresentado na Figura 31 os dois conjuntos de informações estão fortemente relacionados, pois enfatizam o desejo dos colaboradores da empresa por aplicativos para o sistema operacional *Windows phone*®. Como ideia presente no conjunto, está o desejo por aplicativos voltados a relógios inteligentes.

O quarto nodo mais representativo do domínio de análise é o denominado de “*Android app*”. O valor correspondente ao grau é 20. A Figura 33 apresenta o nodo e as ligações diretas com os demais nodos.

Figura 33 – Visualização do nodo *Android app*



Fonte: Autora

Neste nodo existe um desejo dos colaboradores para que a versão do aplicativo *Android*® da Starbucks® seja portado para *FireOS*

Amazon®. Pode-se verificar também o desejo por aplicativos para *Android*® que sejam compatíveis com versões mais antigas deste sistema operacional. No nodo observam-se ainda reclamações por parte dos colaboradores em relação às atualizações disponíveis para o Sistema Operacional *Android*®, que como citado não funcionaram muito bem.

O próximo nodo envolvendo a empresa Starbucks® corresponde ao grupo denominado de “*Drink orders*”. O valor correspondente ao grau do nodo é 17.

Figura 34 – Visualização do nodo *Drink orders*



Fonte: Autora

Neste nodo pode-se observar a presença de ideias como:

- Possibilidade de definir uma loja padrão para realizar encomendas de maneira móvel;
- Aplicativo personalizado para informar o valor nutricional;
- Um cartão de bebida frequente;
- Um botão para repetir a última ordem presente nos registros do usuário;
- Um cartão especial para comprar que, quando utilizado irá lembrar a sua bebida;
- Um aplicativo relacionado à bebida favorita;
- Um aplicativo para realizar uma pré-ordem da bebida e depois buscá-la;
- Possibilidade de solicitar a ordem através do aplicativo móvel.

O sexto e último nodo da análise foi o denominado “*Coffee cups*”. O nodo possui grau 16. Na Figura 35 pode-se visualizar a formação do grupo e suas ideias.

Tabela 3 apresenta o histograma com os 14 termos mais frequentes no terceiro cenário de estudo.

Tabela 3 - Termos mais frequentes no terceiro cenário de estudo

<i>Termo</i>	<i>Número de vezes em que aparecem nos documentos (ideias)</i>
<i>Laptop</i>	32
<i>Alienware</i>	22
<i>Time</i>	21
<i>Computer</i>	20
<i>System</i>	17
<i>Screen</i>	17
<i>Desktop</i>	17
<i>Device</i>	15
<i>Tablet</i>	14
<i>Port</i>	14
<i>Power</i>	14
<i>keyboard</i>	13
<i>Work</i>	13
<i>User</i>	13

Fonte: Autora

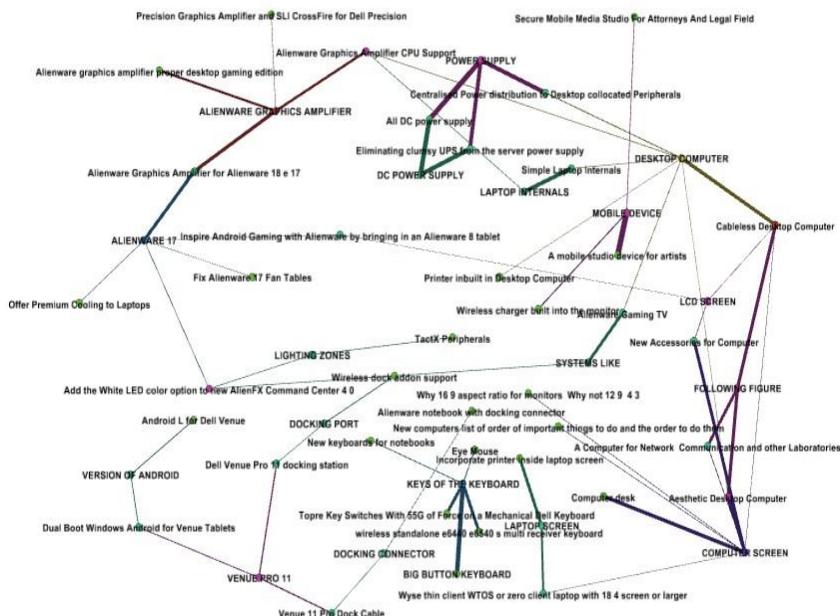
A fim de definir o número de consultas mais adequado ao domínio de análise foi utilizado o índice de modularidade. O objetivo central da modularidade é retornar o valor correspondente ao item de homogeneidade presente no agrupamento e à heterogeneidade presente entre os grupos.

Foram geradas 10 redes com o intuito de identificar qual a melhor formação da rede em relação ao índice de modularidade. Para isto, a primeira rede possuiu como entrada 5 consultas, ou seja, dado o histograma apresentado na Tabela 3, os 5 primeiros termos foram utilizados. A segunda rede foi formada a partir dos 6 primeiros termos do histograma. A terceira rede a partir dos 7 primeiros termos do histograma, e assim por diante, até o total de 14 termos.

A Tabela 4 apresenta os valores para o índice de modularidade obtidos com a aplicação do algoritmo de agrupamento Lingo considerando os termos conforme explicação anterior.

A aplicação do processo de agrupamento gerou um total de 56 nodos e 60 arestas. O valor da modularidade da rede, nas condições apresentadas acima, alcançou um índice de 0,801. Para o Grau médio foi obtido valor correspondente a 2,143; o Diâmetro da rede ficou em 10 e o comprimento médio do caminho foi 4,04. A Figura 37 apresenta o agrupamento geral envolvendo o terceiro cenário de estudo relacionado à empresa Dell®.

Figura 37 – Agrupamento formado a partir do terceiro cenário de estudo



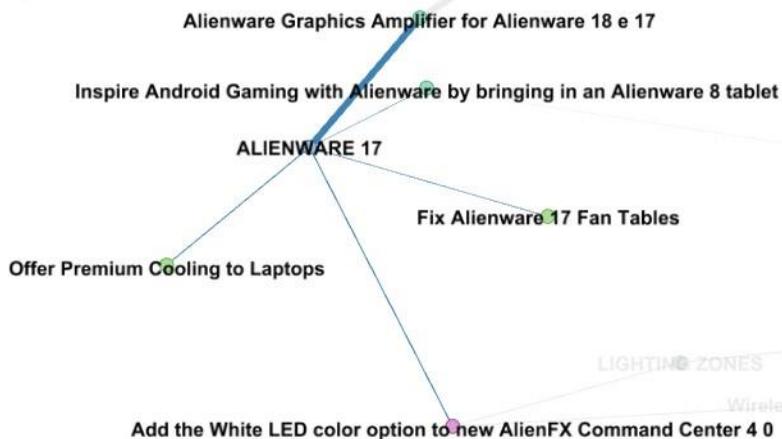
Fonte: Autora

Em virtude do volume de dados optou-se por analisar os nodos com o maior valor correspondente ao item de avaliação “grau”. Para tanto, a Figura 38 apresenta o nodo com o item grau no valor de 8.

Em função das arestas pode-se observar que a ideia relacionada a computador de mesa sem cabo, no caso “*Cableless desktop computer*”, está fortemente relacionada ao nodo denominado “*Desktop computer*”. Outra ideia refere-se à disponibilização de uma impressora embutida a um computador *desktop*. Ocorrem ainda solicitações como a disponibilização padronizada de energia pelo próprio desktop para periféricos por cabos únicos, evitando que estes tenham cabos extras ou que seja necessária a utilização de adaptadores.

A Figura 40 apresenta o terceiro nodo mais representativo no domínio em análise. O grau do nodo atingiu nível 5.

Figura 40 – Visualização do nodo *Alienware 17*

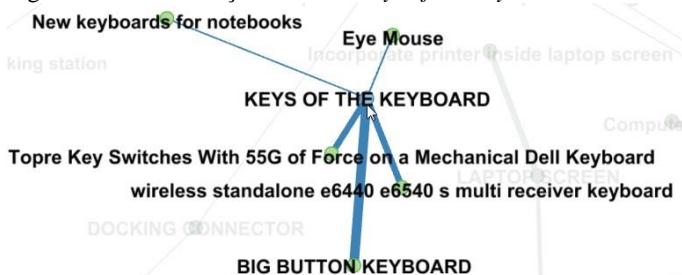


Fonte: Autora

No nodo denominado “*Alienware 17*”, em virtude da espessura da aresta, pode-se observar que os colaboradores da empresa desejam amplificadores de gráficos para o alienware 17 e 18. Alienware são notebooks portáteis da empresa Dell®. Outras ideias relacionadas ao nodo são: “Adicionar a opção de cor LED branco no *AlienFX Command Center 4.0*”, “Jogos para *Android* para o *Alienware*”, “Oferecer resfriamento de alto desempenho para *laptops*” e a “correção dos controles no ventilador do *Alienware 17*”.

A Figura 41 apresenta o quarto nodo mais representativo no domínio em análise. O grau deste nodo atingiu nível 5.

Figura 41 – Visualização do nodo *Keys of the keyboard*



Fonte: Autora

Em função das conexões pode-se observar que o nodo está relacionado com sugestões para teclas referentes ao teclado. As ideias fortemente relacionadas ao nodo sugerem o aumento dos botões do teclado, no caso “*Big button keyboard*”. As outras ideias presentes no nodo referem-se a novos teclados para computadores portáteis, a teclas com 55G de força em um teclado mecânico Dell, e um teclado wireless.

4.2 ANÁLISE DE IMPLEMENTAÇÃO DE IDEIAS DA EMPRESA STARBUCKS®

Objetivando avaliar a efetividade do modelo, realizou-se uma consulta a plataforma *MyStarbucksIdea*® da empresa Starbucks®. A plataforma é responsável pelo gerenciamento das ideias propostas por usuários. Por meio da plataforma é possível acompanhar a avaliação das ideias e as que se encontram em fase de implementação. A pesquisa retornou diversas ideias colocadas em prática pela empresa e que estão diretamente relacionadas com os cenários um e dois, apresentados anteriormente.

No anúncio realizado pelo colaborador 1 da organização, apresentado na Figura 42, pode-se observar o *Starbucks Delivery*, um programa piloto onde os clientes podem ter alimentos e bebidas entregues em suas casas na região de Seattle. Os clientes da organização podem utilizar o aplicativo para iPhone® em bairros de Seattle, tudo de maneira móvel. É possível ainda selecionar a comida e a bebida favorita, personalizá-la, e recebê-la diretamente em casa. A personalização de bebidas e comidas era um forte apelo dos usuários da organização, evidenciado na formação do agrupamento exposto na Figura 30, nodo denominado *Starbucks app*. Além disso, membros da Starbucks® que

utilizam a aplicação ganham recompensas. A solicitação de recompensas por usuários também esteve presente na formação dos agrupamentos, representado na Figura 35.

Figura 42 – Anúncio *Starbucks Delivery in Seattle*

Starbucks Delivery in Seattle

Dec 02, 2015
by Colaborador 1

RSS



Starbucks delivery is one of the **most popular topics of all time** on My Starbucks Idea. We've always loved the idea, and are very excited to announce the arrival of Starbucks Delivery, a pilot program where customers can have food and drinks delivered to them within designated areas in Seattle. Through a collaboration with Postmates, a leading on-demand delivery service, Starbucks customers using the Starbucks app for iPhone in designated Seattle neighborhoods, will now be able to choose "Delivery" as an option within Mobile Order & Pay. Simply select your favorite food and beverage, customize it just the way you like, and have it delivered directly to you. Plus, My Starbucks Rewards members will earn stars with each delivery order.

Fonte: Starbucks® (2016)

Na Figura 43 apresenta-se uma alternativa a iluminação ambiente da organização. A especialista e ambientalista, colaboradora 2 da organização Starbucks®, anuncia a utilização de Lâmpadas LED para tornar mais eficiente o consumo de energia. Uma parceria com General Electric™ permitiu a organização produzir lâmpadas LED eficientes. Também foram introduzidas lâmpadas LED estilo *Edison*. Estas lâmpadas personalizadas utilizam menos energia.

Figura 43 – Anúncio *Saving energy with LED Lighting*

Saving energy with LED Lighting

May 11, 2015
by Colaborador 2



Since 2008, you have shared a lot of ideas about how we can make our stores more "green," more comfortable, and more energy efficient. I work on the Facilities Environmental Performance Management Team, and as an environmentalist myself, I am proud to share some of the great strides we've made with energy efficient lighting in our stores. But first, let me share an email from my colleague's Mom: "This past week I was at my Starbucks getting coffee and the physical setting was so friendly, smelled great, and of course people genuinely friendly. I asked what had changed to make this store so comfortable....changed the lighting!!!!!! Hope your day is productive or relaxed ... **more** »

1 comments

Launched, Atmosphere & Locations, Social Responsibility

Fonte: Starbucks® (2016)

O anúncio presente na Figura 43 demonstra o apelo dos usuários da organização por uma organização mais verde, sustentável e eficiente em energia. Sugestões de ideias para tornar a organização mais verde foram apresentadas na Figura 35. Também na Figura 35 ocorre o apelo por locais que produzam energia através de reciclagem.

Na Figura 44 o colaborador 3 da organização anuncia uma atualização para o aplicativo que permite que usuários dos EUA, Canadá e Reino Unido realizem sincronizações com Apple Watch™. Com o aplicativo Starbucks® para iPhone® v3.2.5 instalado e sincronizado com o Apple Watch™, o usuário poderá ver o seu saldo, acompanhar suas recompensas, encontrar lojas Starbucks® nas proximidades, acompanhar saldos e detalhes do *Starbucks Card* e pagar suas ordens através do Apple Watch™ usando o Passbook®.

Figura 44 – Anúncio *Starbucks app + Apple Watch*

Starbucks app + Apple Watch

Apr 24, 2015
by Colaborador3

RSS



Over the years, your ideas have inspired many new features and functions for the Starbucks® app for iPhone® including: Pick of the Week, Mobile Order & Pay, shake-to-pay, digital tipping, Card reload with Apple Pay™, and tracking Stars and earned rewards.

Today we released an update to our app that allows app users in the U.S., Canada, and the U.K. to synch with Apple Watch™.

With the Starbucks app for iPhone® v3.2.5 installed and synced with Apple Watch, you'll be able to:

- View your My Starbucks Rewards® Star balance and track earned rewards
- Find nearby Starbucks® stores, including those with a drive-thru or Starbucks Reserve® coffees
- See all of your Starbucks Cards balances and card details
- Pay with your Apple Watch in participating Starbucks stores using Passbook®

Fonte: Starbucks® (2016)

A presença de ideias relacionadas a relógios inteligentes foi apresentada na Figura 30, onde se ressalta o apelo por parte dos usuários da organização por novos aplicativos para relógios inteligentes, com a formação do agrupamento denominado *Starbucks app*.

Na Figura 45 o colaborador 4 anuncia a criação de um programa de recompensas para os usuários membros. O usuário ao pagar com seu cartão *Starbucks Rewards* ou via aplicativo receberá descontos especiais em diferentes datas.

Figura 45 – Anúncio *Happy Member Mondays*

Happy Member Mondays
 March 06, 2015
 by Colaborador4



My Starbucks Rewards® members love getting rewards, and we love giving them. So when we saw all the ideas about expanding perks for rewards members, we just couldn't help ourselves. Starting March 9, when you pay with your registered Starbucks Card or app, My Starbucks Rewards members can enjoy special Happy Monday offers from 2-5 pm every Monday in March at participating stores. March 9 – 30% off food, drinks, and more March 16 – Two tea beverages for the price of one March 23 – Free pastry when you buy any handcrafted drink March 30 – Frappuccino Happy Hour with 50% off Frappuccino blended beverages Cue happy song.

17 comments

Coming Soon, Launched, Coffee & Espresso Drinks, Tea & Other Drinks, Merchandise & Music, Starbucks Card

Fonte: Starbucks® (2016)

A presença de ideias relacionadas a programas de recompensas foi evidenciada na formação do agrupamento denominado *Coffee cups*, Figura 35.

Na Figura 46 ocorre o anúncio de atualizações para aplicativos com o sistema operacional Android™, disponível no *Google Play Store*. Dentre as funcionalidades do aplicativo está uma nova caixa de entrada. Com a nova caixa de entrada, o usuário será informado sobre eventos e ofertas exclusivas, e também terá acesso antecipado para selecionar produtos direito em sua aplicação.

Figura 46 – Anúncio *Android App Update*

Android App Update
 January 08, 2015
 by Colaborador5



Your ideas and feedback about our app has been helpful and inspiring. We've heard from a number of you about updating the Android™ app, and we're excited to announce that we have! Our latest version of the Starbucks® app for Android™ is available now in the Google Play Store and features a new inbox and Android Wear support (USA only). With the new inbox, we'll let you know about exclusive events and offers, early access to select products — right in your app. The new release also brings wearables support to the app, so you can get tip and inbox notifications right on your Android Wear device. This release expands the features of the new Starbucks® app ... [more »](#)

7 comments

Launched, Starbucks Card

Fonte: Starbucks® (2016)

Ideias relacionadas a atualizações para aplicativos com o sistema operacional Android™ foram evidenciadas na Figura 30 e Figura 33.

A Figura 47 apresenta uma atualização para aplicativos com o sistema operacional iPhone®, que inclui duas características importantes como o menu móvel e o *Order & Pay*. No menu do celular, o usuário pode navegar rapidamente e pesquisar por itens novos ou favoritos, bem como, obter informações nutricionais (quando disponível). A questão nutricional era um forte apelo apresentado na formação dos agrupamentos, seção anterior, Figura 25 e Figura 26. A pesquisa por itens favoritos também foi evidenciado nos cenários acima, Figura 30 e Figura 34. O menu possibilita destacar os produtos que estão disponíveis a nível nacional. Na opção *Order & Pay* é possível encomendar e pagar a ordem diretamente do telefone, solicitação esta apresentada na Figura 30. E existe ainda a possibilidade de personalizar totalmente as bebidas via aplicativo, ideia presente também na Figura 30. Este último ponto também era um forte apelo constatado na formação dos agrupamentos.

Figura 47 – Anúncio *Mobile Ordering Coming to Portland, OR and Menu Returning Through the Updated Starbucks® app for iPhone®*

Mobile Ordering Coming to Portland, OR and Menu Returning Through the Updated Starbucks® app for iPhone®

Dec 04, 2014 RSS

by Colaborador3



We love hearing feedback and **ideas** from our customers and strive to make sure we're providing features that are both relevant and convenient for you. Today we're excited to release an update to the Starbucks® app for iPhone® that includes two important features you've told us you want: mobile menu and Mobile Order & Pay.

- **Mobile Menu:** It's back – a version of our Starbucks menu has returned to the app. Now you can quickly browse and search the mobile menu to find and try new or favorite items as well as get nutritional information (when available). Right now, the menu highlights products that are available nationally and we have plans to expand the mobile menu further so that it contains all of your favorite Starbucks items.
- **Mobile Order & Pay:** Available first in participating stores in Portland, Oregon, the new feature of the Starbucks® app for iPhone® allows you to order and pay for your Starbucks order directly from your phone. Just as you would when ordering with a barista, you can also fully customize your drinks right from the app.

Fonte: Starbucks® (2016)

Com as constatações acima, pode-se observar que a organização trabalha na perspectiva de atender as ideias de seus usuários, objetivando fornecer recursos relevantes e convenientes para os mesmos. A análise promovida pelo conjunto de especialistas da organização evidencia as descobertas efetuadas pelo modelo.

4.3 CONSIDERAÇÕES FINAIS

Este capítulo teve como intuito demonstrar a viabilidade do modelo proposto. Neste sentido, inicialmente foi desenvolvido um protótipo dividido basicamente em cinco fases: a primeira fase consiste na estruturação do conjunto de ideias por meio da criação de documentos no formato XML, a fim de facilitar a manipulação através de meios computacionais.

A segunda fase é responsável pela indexação desta informação, por meio da leitura dos documentos no formato XML através de uma biblioteca de indexação de documentos, neste caso, Lucene Apache®.

A terceira fase consiste no preenchimento da ontologia de domínio. A quarta fase é representada pelo processamento das ideias por meio da realização do processo de agrupamento, utilizando como algoritmo o Lingo presente no projeto Carrot².

A quinta e última fase envolve a explicitação e avaliação do conhecimento por meio de ferramentas de visualização de grafos no padrão Graphml;

O protótipo desenvolvido tem como finalidade a descoberta de novas sugestões de ideias (utilizando o processo de agrupamento), a população de uma ontologia de domínio e a visualização das informações através de redes. Por meio do protótipo foram realizadas análises considerando conjuntos de ideias coletadas a partir das empresas Starbucks® e Dell®. Os resultados obtidos permitiram explicitar visualmente os relacionamentos presentes entre as ideias coletadas. Por fim, a identificação de ideias implementadas pela empresa Starbucks® trouxe comprovação das descobertas realizadas pelo modelo.

5 CONSIDERAÇÕES FINAIS

Neste capítulo serão apresentadas as considerações finais obtidas com o desenvolvimento do trabalho, bem como as contribuições geradas pela aplicação da pesquisa e as perspectivas de trabalhos futuros.

5.1 CONCLUSÕES

O presente trabalho apresentou um modelo baseado em ontologia e análise de agrupamento para auxiliar no processo de tomada de decisão no domínio de Gestão de Ideias. A fim de atingir os objetivos, uma ontologia de domínio foi desenvolvida como forma de representação do conhecimento e para fornecer apoio à representação semântica das ideias. A representação semântica contribui no processo de agrupamento de ideias em categorias, na busca e na avaliação das ideias.

Através da semântica é possível determinar as regras para combinar ideias e seus relacionamentos, promovendo a representação dos relacionamentos. O uso de ontologias permite o raciocínio, a análise e a atuação sobre o conhecimento armazenado.

O modelo considera também técnicas de análise multivalorada. Essas técnicas possibilitam a avaliação de um conjunto de características em relação a uma população em análise. Entre as técnicas de análise multivalorada com alta empregabilidade está a análise de agrupamento. A análise de agrupamento consiste em uma técnica que permite agrupar objetos com base nas semelhanças de cada observação. A ideia central na formação dos agrupamentos é manter a homogeneidade dentro do grupo formado e a heterogeneidade entre os grupos. Com a utilização da técnica de agrupamento é possível reduzir as informações de um conjunto de “*n*” indivíduos para informações de um novo conjunto de “*k*” grupos.

Partindo deste pressuposto é plausível observar que através da análise de agrupamento é possível evidenciar padrões e tendências com relação às ideias analisadas em um domínio. Analisar individualmente uma ideia pode não representar um potencial de implementação, porém ao analisar um conjunto com outras ideias, pode-se evidenciar uma possível tendência ou demanda de mercado.

Agrupamentos permitem que usuários interpretem de uma melhor maneira as ideias coletadas na organização. Com isto, o processo de tomada de decisão é facilitado, auxiliando especialistas a decidirem

sobre quais ideias são passíveis de serem implementadas e requerem futuros investimentos.

O modelo proposto foi aplicado por meio de cenários de estudo envolvendo as empresas Starbucks® e Dell®. Observou-se nos dois primeiros cenários envolvendo a empresa Starbucks® que existe um forte apelo para o desenvolvimento de aplicativos *mobile* que utilizam o sistema operacional *Windows Phone*®, por aplicativos que forneçam informações nutricionais ou ainda, calculadoras nutricionais, e por aplicativos relacionados ao domínio de bebidas.

No terceiro cenário de estudo envolvendo a empresa Dell®, constatou-se que existe um apelo por questões voltadas à área gráfica do *laptop* modelo Alienware®, por novos acessórios, impressoras embutidas em desktops, resfriamento de laptops e relacionados ao domínio de jogos. Analisando o conjunto de ideias coletadas observou-se ainda que os colaboradores da empresa Dell® utilizam a plataforma denominada *ideastorm* como uma forma de obter suporte ou ainda fazer reclamações.

Através das consultas realizadas ao site da empresa Starbucks®, pode-se observar que a organização trabalha na perspectiva de atender as ideias de seus usuários com o intuito de satisfazer suas necessidades, fornecendo recursos relevantes e convenientes. A análise das ideias anunciadas no site da organização e colocadas em prática comprova as descobertas efetuadas pelo modelo.

As informações presentes na base de conhecimento permitem a formação e apresentação de grupos de ideias, possibilitando que especialistas tenham um ferramental que pode reduzir o tempo dispendido na análise de tendências e demandas apontadas por clientes e colaboradores.

5.2 PERSPECTIVAS DE TRABALHOS FUTUROS

Como trabalhos futuros, vislumbram-se a evolução do algoritmo responsável pelo processo de agrupamento a partir de um conjunto de ideias. No modelo atual, as ideias são agrupadas a partir da geração de um histograma que funciona como um meio para identificar as sementes (termos mais frequentes) a serem utilizadas no processo.

Outros algoritmos de agrupamento podem ser aplicados com o objetivo de analisar qual a melhor abordagem para lidar com informações textuais. Além disso, o aprimoramento das análises dos mapas produzidos pelo processo de agrupamento através da análise de

redes sociais pode tornar-se algo relevante para o domínio de estudo voltado às ideias.

A Gestão de Críticas recebidas através de colaboradores ou consumidores mostra-se como um cenário interessante para o modelo, possibilitando o manejo de críticas para ressignificar os discursos de gestão, e como um método alternativo para analisar e explorar impactos referentes ao desempenho da organização no mercado junto ao seu público alvo.

Outra possibilidade é considerar a dimensão tempo no agrupamento e avaliação de ideias. Na análise de agrupamento baseada em termos, a frequência pode influenciar na formação da rede. Neste sentido, fatiar as ideias por períodos de tempo antes de realizar o processo de agrupamento pode promover análises mais consistentes, pois irá reduzir o número de ideias para um período particular de interesse do analista.

Sugere-se uma maior exploração das tecnologias presentes na Web Semântica, ao qual representam um grande potencial para área, possibilitando organizar as estruturas de dados a partir de conceitos e relacionamentos, assim como, realizar inferências a partir dessas estruturas.

Por fim, com o objetivo de atingir resultados mais consistentes no processo de identificação de ideias passíveis de implementação e investimentos, bem como, auxiliar na visualização destas informações, poderia ocorrer o desenvolvimento de um sistema de recomendação, que com base em algumas regras sugeridas pelo especialista de domínio e, dado uma pré-classificação do colaborador da empresa, indicar qual a melhor ideia a ser implementada.

REFERÊNCIAS

- AAGAARD, A.; GERTSEN, F. Supporting radical front end innovation: perceived key factors of pharmaceutical innovation. **Creativity and Innovation Management**, Malden, vol. 20, n°. 4, p. 330-346, 2011.
- ADAMS, R., BESSANT, J., PHELPS, R. Innovation management measurement: a review. **International Journal of Management Reviews**, vol. 8, n°.1, p. 21-47, 2006.
- ALJABER B. *et al.* Document *clustering* of scientific texts using citation contexts. **Information Retrieval Journal**, Springer Science+Business Media, 2009.
- ALMEIDA, Mauricio B.; BAX, Marcello P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ci. Inf.**, Brasília, v. 32, n. 3, Dec. 2003.
- ANDRIENKO, G., ANDRIENKO, N. Interactive *cluster* analysis of diverse types of spatiotemporal data. **Acm Sigkdd Explorations Newsletter**, New York, v. 11, n° 2, p. 19-28, 2009.
- ARTZ, K. W., *et al.* A longitudinal study of the impact of R&D, patents, and product innovation on firm performance. **Journal of Product Innovation Management**, vol. 27, n°.5, p. 725-740, 2010.
- BARBIERI, J. C.; ÁLVARES, A. C. T.; CAJAZEIRA, J. E. R. **Gestão de Ideias para inovação contínua**. Porto Alegre: Bookman, 2009.
- BAREGHEH, A.; ROWLEY, J.; SAMBROOK, S. Towards a multidisciplinary definition of innovation. **Management Decision**, United Kingdom, v. 47, n .8, p.1323-1339, 2009.
- BATZIAS, F. A.; SIONTOROU, C. G. Creating a specific domain ontology for supporting R&D in the science-based sector – The case of biosensors. **Expert Systems with Applications**, v. 39, n. 11, p. 9994-10015, 2012. Elsevier Ltd.
- BAUTZER, Deise. **Inovação: repensando as organizações**. São Paulo: Atlas, 2009.

BERRY, M. J. A.; LINOFF G. **Data mining techniques – for marketing, sales and customer support**. John Wiley & Sons, New York, 1997.

BESSANT, J. *et al.* Managing innovation beyond the steady state. **Technovation**, Amsterdam, vol. 25, n°. 12, p. 1366-1376, 2005.

BJÖRK, J.; BOCCARDELLI, P.; MAGNUSSON, M. G. Ideation capabilities for continuous innovation. **Creativity & Innovation Management**, Malden, v. 19, n. 4, p. 385-396, 2010.

BLONDEL, Vincent D. *et al.* Fast unfolding of communities in large networks. **Journal of Statistical Mechanics: Theory and Experiment**, v. 2008, n. 10, p. P10008, 2008.

BORST, W. N. **Construction of engineering ontologies for knowledge sharing and reuse**. Universiteit Twente, 1997.

BOTHOS, E.; APOSTOLOU, D.; MENTZAS, G. Collective intelligence with web-based information aggregation markets: The role of market facilitation in idea management. **Expert Systems with Applications**, Amsterdam, vol. 39, n°. 1, p. 1333-1345, 2012.

BREM, A.; VOIGT, K. I. Innovation management in emerging technology ventures: the concept of an integrated idea management. **International Journal of Technology, Policy and Management**, Olney, v. 7, n. 3, p. 304-321, 2007.

BURRELL, G.; MORGAN, G. **Sociological Paradigms and Organizational Analysis**. Heineman: London, 1979.

CAO, L.; ZHAO, Y.; ZHANG, H.; LUO, D.; ZHANG, C.; PARK, K. Flexible Frameworks for Actionable Knowledge Discovery. **IEEE Transactions on Knowledge and Data Engineering**, vol.22, n. 9, p. 1299-1312, Sept. 2010.

CARBONE, F.; CONTRERAS, J.; HERNÁNDEZ, J. Z.; GOMEZ-PEREZ, J. M. Open Innovation in an Enterprise 3.0 framework: Three 208 case studies. **Expert Systems with Applications**, v.39, n. 10 February, p. 8929-8939, 2012.

CHANG, S. L.; CHEN, C. Y.; WEY, S. C. Conceptualizing, assessing, and managing front-end fuzziness in innovation/NPD projects. **R&D Management**, Malden, v. 37, n. 5, p. 469-478, nov. 2007.

CLARK, K. B.; WHEELWRIGHT, S. C. **Managing new product and process development**: text and cases. New York: The Free Press, 1993, p. 293.

CONRY-MURRAY. (2010). Can Enterprise Social Networking Pay Off? Internet Evolution. [Online]. Disponível em: http://www.internetevolution.com/document.asp?doc_id=173854&page_number=2; Acesso em: Jan, 2015.

CONVERTINO, Gregorio; GRASSO, Antonietta; MICHELIS, Giorgio De; MILLEN, David R.; CHI, Ed H. Clorg: collective intelligence in organizations tools and studies. **In Proceedings of the 16th ACM international conference on Supporting group work (GROUP '10)**. ACM, New York, NY, USA, 355-358. 2010.

COOK, K. S.; EMERSON, R. M.; GILLMORE, M. R. The Distribution of Power in Exchange Networks: Theory and Experimental Results. **The American Journal of Sociology**, v. 89, n. 2, p. 275-305, 1983.

COOPER, R. G.; EDGETT, S. J. **Successful product innovation**: a collection of our best. Charleston: Booksurge, 2009.

COOPER, R. G.; EDGETT, S. J.; KLEINSCHMIDT, E. J. **Portfolio management for new products**. 2. ed. New York: Basic Books, 2001.

COOPER, R. G.; EDGETT, S. J.; KLEINSCHMIDT, E. J.; Optimizing the Stage-Gate process: What best practice companies are doing – Part I. **Research Technology Management**, v. 45, n. 5, 2002.

COOPER, Robert G. Stage-gate systems: A new tool for managing new products, **Business Horizons**, vol. 33, n. 3, p. 44-54, 1990.

CORMACK, R. M. A review of classifications. **Journal of the Royal Statistical Society**, vol. 134, n° 3, p. 321-367, 1971.

CRAWFORD, M.; BENEDETTO, A. D. **New products management**. Irwin: McGraw-Hill, 2006.

CUPANI, Alberto. **Filosofia da Tecnologia: um convite**. Florianópolis: Editora da UFSC, 2011.

DESCHAMPS, J. P.; NAYAK, P. R.; LITTLE, A. D. **Wachstum und Gewinn durch offensive Produktstrategien**, Frankfurt/Main: Campus Verlag, 1996.

DOROW, Patrícia Fernanda. **O Processo de Geração de Ideias para Inovação: Estudo de Caso em uma Empresa Náutica**. 2013. 158 f. Dissertação (Mestrado) - Programa de Pós-graduação em Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina, Florianópolis, 2013.

EASLEY, D.; KLEINBERG, J. **Networks, crowds, and markets**. Cambridge Univ Press. 2010.

EBECKEN, Nelson F. F.; LOPES, Maria Celia S.; COSTA, Myrian C. A. Mineração de texto. In: REZENDE, Solange O. (Coord.). **Sistemas inteligentes: fundamentos e aplicações**. São Paulo: Manole, 2005.

EGC. **Áreas de Concentração**. 2016. Disponível em: <<http://www.egc.ufsc.br/pos-graduacao/programa/areas-de-concentracao/>>. Acesso em: 16 jan. 2016.

EGC. **Linhas de Pesquisa**. Disponível em: <<http://www.egc.ufsc.br/index.php/pesquisas/linhas-de-pesquisa>>. Acesso em: 30 out. 2014.

ENDESLEY, S. **Innovation in action: a practical guide for healthcare teams**. London: BMJ Books, 2010.

EVERITT, Brian S. **A handbook of statistical analyses using S-Plus**. CRC Press, 2001.

FAYYAD, U. *et al.* **From Data Mining to Knowledge Discovery in Databases**. AAAI/MIT Press, 1997.

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. From Data Mining to Knowledge Discovery in Databases. **AI Magazine**, n. 17, p. 37-54. 1996.

FELDMAN, R. *et al.* A domain independent environment for creating information extraction modules, In: **Proceedings of the tenth international conference on information and knowledge management**, p 581-588, ACM Press. 2001

FELDMAN, R.; DAGAN, I. Knowledge discovery in textual databases (KDT). In: **Knowledge Discovery and Data Mining**, p. 112–117, 1995.

FENN, Jackie; LEHONG, Hung. Hype cycle for emerging technologies, 2011. **Gartner**, July, 2011.

FERNANDES, R. P.; GROSSE, I. R.; KRISHNAMURTY, S.; WITHERELL, P.; WILEDEN, J. C. Semantic methods supporting engineering design innovation. **Advanced Engineering Informatics**, v. 25, n. 2, p. 185-192, 2011.

FLYNN, M. *et al.* Idea management for organizational innovation. **International Journal of Innovation Management**, Washington, v. 7, n. 5, p. 417-442, 2003.

FRAWLEY, W. J.; PIATETSKY-SHAPIRO, G.; MATHEUS, C. J. Knowledge discovery in databases: an overview. **AI Mag**, vol. 13, n. 3, p. 57-70, 1992.

FREEMAN, Linton C. Centrality in social networks conceptual clarification. **Social networks**, v. 1, n. 3, p. 215-239, 1979.

FREITAS, Frederico Luiz G. de. Ontologias e a web semântica. In: **Congresso da Sociedade Brasileira de Computação**, 28., 2003, Campinas. Anais... Campinas: SBC, 2003. p. 1-52.

FUNG, Glenn. **A comprehensive overview of basic clustering algorithms**. 2001.

GAMA, J., *et al.* **Extração de Conhecimento de Dados - Data Mining**. Lisboa, Edições Sílabo, LDA, 2012.

GAMBATESE, J. A.; HALLOWELL, M. Enabling and measuring innovation in the construction industry. **Construction Management and Economics**, vol. 29 n° 6, 2011.

GARDNER, S.P. Drug Discovery Today: Technologies. **Knowledge Management**, v. 2 n. 3, 2005.

GAYNOR, G.H. Innovator: what does it take to be one? **IEEE Antennas and Propagation Magazine**, vol. 43, n°. 3, p. 126–130, 2001.

GEPHI. **Gephi Makes Graphs Handy**. Disponível em: <<http://gephi.github.io/>>. Acesso em: 30 jun. 2015.

GI2MO. **GI2MO Semantically empowered idea management**. Disponível em: <<http://www.gi2mo.org/>>. Acesso em: 10 jun. 2014.

GIBSON, R.; SKARZYNSKY, P. **Inovação: prioridade nº 1: o caminho para transformação nas organizações**. Tradução: Alessandra Mussi Araujo, Rio de Janeiro: Elsevier, 2008.

GIL, Antônio Carlos. **Métodos e técnicas de pesquisa social**. São Paulo: Atlas, 1999.

GIRVAN, Michelle; NEWMAN, Mark EJ. Community structure in social and biological networks. **Proceedings of the National Academy of Sciences**, v. 99, n. 12, p. 7821-7826, 2002.

GÓMEZ-PÉREZ, A. Ontological engineering: A state of the art. Expert Update: **Knowledge Based Systems and Applied Artificial Intelligence**, v. 2, n. 3, p. 33-43, 1999.

GONÇALVES, Alexandre Leopoldo. **Um modelo de descoberta de conhecimento baseado na correlação de elementos textuais e expansão vetorial aplicado à engenharia e gestão do conhecimento**. Florianópolis, SC, 2006. 196 f. Tese (Doutorado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Engenharia de Produção.

GOWER, J. C.; LEGENDRE, P. Metric e euclidean properties of dissimilarity coefficients. **Journal of Clasification**, v. 3, p. 5 - 48, 1986.

GRUBER, T. R. Toward principles for the design of ontologies used for knowledge sharing? **International journal of human-computer studies**, v. 43, n. 5, p. 907-928, 1995.

GRUBER, T.. **What is an Ontology?** 1993. Disponível em: <<http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>>. Acesso em: Jan. 2015.

GUARINO, N. Formal Ontology in Information Systems: **Proceedings of the 1st International Conference** June 6-8, 1998, Trento, Italy. 1998.

HAIR, J. F. *et al.* **Multivariate Data Analysis**. 7. ed. Pearson Prentice Hal, 2010. 593 p.

HALKIDI, Maria; BATISTAKIS, Yannis; VAZIRGIANNIS, Michalis. On clustering validation techniques. **Journal of Intelligent Information Systems**, vol. 17, p. 107-145, 2001.

HAN, J.; KAMBER, M. **Data Mining: Concepts and Techniques**. 1^a ed. Nova York: Morgan Kaufmann Publishers, 2001.

HEARST, Marti A. Untangling text data mining. **In: Association For Computational Linguistics on Computational Linguistics**, 37., 1999, Stroudsburg. Proceedings Stroudsburg: Association For Computational Linguistics, p. 3 – 10, 1999.

HÜSIG, Stefan; KOHN, Stefan. “Open CAI 2.0”–Computer Aided Innovation in the era of open innovation and Web 2.0. **Computers in Industry**, v. 62, n. 4, p. 407-413, 2011.

IQBAL, Adnan. Creativity and innovation in Saudi Arabia: An overview. **Innovation: Management. Policy & Practice**, vol. 13, n^o. 3, p. 376-390, 2011.

JACKSON, A. A.; SOMERS, K. M.; HARVERY, H. H. Similarity coefficients: measures for co-occurrence e association or simply measures of occurrence. **American Naturalist**, v.13, p. 436 - 453, 1989.

JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. Data *clustering*: a review. **Acm Computing Surveys**, v. 31 n^o 3, p. 264-323, 1999.

JASPER, R; USCHOLD, M. A Framework for Understanding and Classifying Ontology Applications. **Proceedings of the IJCAI-99**

Workshop on Ontologies and Problem Solving Method (KRR5), Stockholm, Sweden, 1999.

JENSEN, P.H., WEBSTER, E. Another look at the relationship between innovation proxies. **Australian Economic Papers**, vol. 48, n°. 3, p. 252–269, 2009.

JSON. **Introdução ao JSON**. Disponível em: <http://www.json.org/json-pt.html>. Acesso: 16 jan. 2016.

JUNG, J. J. Semantic business process integration based on ontology alignment. **Expert Systems with Applications**, v. 36, n. 8, p. 11013–11020, 2009

KAINULAINEN, Jukka. **Clustering algorithms: basics and visualization**. Helsinki University of Technology, Laboratory of Computer and Information Science, 2002.

KOBASHI, Nair Yumiko. Fundamentos semânticos e pragmáticos da construção de instrumentos de representação de informação. **DataGramaZero - Revista de Ciência da Informação**, v. 8, n. 6, 2007.

KOEN, Peter A. *et al.* **Fuzzy front end: Effective methods, tools, and techniques**. Wiley, New York, NY, 2002.

KONCHADY, Manu. **Text mining application programming**. Massachusetts: Charles River Media, 2006.

KUECHLER, W. L. Business applications of unstructured text. **Communications of ACM**, vol. 50, n. 10, p. 86-93, 2007.

KUMAR, V. S.; GRESS, C. L. Z.; HADWIN, A. F.; WINNE, P. H. Assessing process in CSCL: An ontological approach. **Computers in Human Behavior**, v. 26, n. 5, p. 825–834, 2010.

LATTIN, J. M.; DOUGLAS C.; PAUL E. G. **Análise de dados multivariados**. São Paulo: Cengage Learning, 2011. 455 p.

LATTIN, J.; CARROLL, J. D.; GREEN, P. E. **Analyzing Multivariate Data**. Cole: Thomson Books, 2003.

LE, J.; Portfolio management for projects. In: **Engineering Management Conference**, out. 2004, Singapura. IEEE International. v.3, n.18-21, 2004. p. 1013-1017.

LI, Xingsen; LI, Liping; CHEN, Zhengxin. Toward extenics-based innovation model on intelligent knowledge management. **Annals of Data Science**, v. 1, n. 1, p. 127-148, 2014.

LINDER, Stefan. Fostering strategic renewal: monetary incentives, merit-based promotions, and engagement in autonomous strategic action. **Journal of Management Control**, p. 1-30, 2015.

LULA, P.; PALIWODA-PEKOSZ, G.. An Ontology-Based *Cluster* Analysis Framework. **Proceedings Of The First International Workshop On Ontology-supported Business Intelligence**, New York, p.1-6, 2008.

MACQUEEN, J. B. Some methods for classification and analysis of multivariate observations. **Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability**, vol. 1: Statistics, p. 281-297, University of California Press, Berkeley, Calif., 1967.

MAEDCHE, Alexander. **Ontology Learning for the Semantic Web**, 2002.

MAGNUSSON, Peter R.; NETZ, Johan; WÄSTLUND, Erik. Exploring holistic intuitive idea screening in the light of formal criteria, **Technovation**, vol. 34, n. 5-6, May-June 2014, Pages 315-326, 2014.

MAHROUM, Sami; AL-SALEH, Yasser, Towards a functional framework for measuring national innovation efficacy, **Technovation**, vol. 33, n°. 10-11, 2013, p. 320-332, ISSN 0166-4972.

MAIA, L. C. G.; ROCHA, R. S.. Uso de sintagmas nominais na classificação automática de documentos eletrônicos. **Perspectivas em Ciência da Informação**, v. 15, n. 1, p.154-172, 2010.

MANNING, C. D.; SCHÜTZE, H. **Foundations of Statistical Natural Language Processing**. MIT Press, 2003.

MARZANO, Stefano. People as a source of breakthrough innovation. **Design Management Review**, v. 16, n. 2, p. 23- 29, 2005.

MCADAM, R.; MCCLELLAND, J. Individual and team-based idea generation within innovation management: organizational and research agendas. **European Journal of Innovation Management**, United Kingdom, v. 5, n. 2, p. 86-97, 2002.

MCGRATH, M. **Next generation product development**: how to increase productivity, cut costs, and reduce cycle times. New York: McGraw-Hill, 2004.

MIKA, P.; AKKERMANS, H. Towards a new synthesis of ontology technology and knowledge management. **The Knowledge Engineering Review**, v. 19, n. 4, p. 317-345, 2005.

MOOS, B.; WAGNER, H.-T.; BEIMBORN, D.; WEITZEL, T., "The Role of Innovation Governance and Knowledge Management for Innovation Success, **44th Hawaii International Conference on System Sciences (HICSS)**, Jan. 2011.

MOWELL, J. M.; SHEA, C. M. Individual differences, environmental scanning, innovation framing, and champion behavior: key predictors of project performance. **The Journal of Product Innovation Management**, v. 18, p. 15-27, 2001.

NECHES, R.; FIKES, R.; FININ, T.; GRUBER, T.; PATIL, R.; SENATOR, T.; SWARTOUT, W. Enabling Technology for Knowledge Sharing. In **AI Magazine**, Fall, 1991.

NEWELL, A. The Knowledge Level. (Presidential address, AAAI-80). **AI magazine**, p. 1-20, 1981.

NEWMAN, M. E. J. Modularity and community structure in networks. **Proc.Natl.Acad.Sci.Usa**, 103, 2006.

NEWMAN, M. E. J.; GIRVAN, M. Finding and evaluating community structure in networks. **Physical Review E, American Physical Society**, v. 69, n. 2, 2004.

NEWMAN, M. E. The structure and function of complex networks. **SIAM review** vol. 45, n° 2, p. 167-256, 2003a.

NEWMAN, M. Fast algorithm for detecting community structure in networks. **Physical Review E**, v. 69, set. 2003b.

NOACK, T.; SCHMITT, I. Monitoring mobile cyber-physical systems by means of a knowledge discovery cycle. **IEEE Seventh International Conference on Research Challenges in Information Science (RCIS)**, p. 1-12, May 2013.

NOBELIUS, D.; TRYGG, L. Stop chasing the front end process: management of the early phases in product development projects. **International Journal of Project Management**, Amsterdam, v. 20, n. 5, p. 331-3340, 2002.

OKE, A., MUNSHI, N., WALUMBWA, F.O. The influence of leadership on innovation processes and activities. **Organizational Dynamics**, vol. 38, n°. 1, p. 64-72, 2009.

OSIŃSKI, Stanisław. **An algorithm for clustering of web search results**. Tese de Doutorado. Poznań University of Technology, Poland. 2003.

OSIŃSKI, Stanisław; STEFANOWSKI, Jerzy; WEISS, Dawid. Lingo: Search results *clustering* algorithm based on singular value decomposition. **In: Intelligent information processing and web mining**. Springer Berlin Heidelberg, 2004. p. 359-368.

PALAZZO, M. D. O.; LOH, S.; AMARAL, L. A.; WIVES, L. K. Descoberta de conhecimento em textos através da análise de sequências temporais. In: **Workshop em Algoritmos e Aplicações de Mineração de Dados - WAAMD, SBBD**: Sociedade Brasileira de Computação, volume II, p. 49-56, Florianópolis, 2006.

PAUKKERI, Mari-Sanna; KOTRO, Tanja. Framework for analyzing and clustering short message database of ideas. **Proceedings of I-KNOW '09 and I-SEMANTICS '09**, Graz, Austria 2009.

PETTEY, C.; STEVENS, H. Gartner's 2009 hype cycle special report evaluates maturity of 1,650 technologies. **Tech. rep.** 2009.

POLI, Roberto; OBRST, Leo. The interplay between ontology as categorial analysis and ontology as technology. In: **Theory and applications of ontology: Computer applications**. Springer Netherlands, 2010. p. 1-26.

POVEDA, G.; WESTERSKI, A.; IGLESIAS, C. A. Application of semantic search in Idea Management Systems. **International Conference for Internet Technology And Secured Transactions**, 2012, vol., no., p.230 - 236, 10-12 Dec. 2012

QUINTANE, E. *et al.* Innovation as knowledge-based outcome. **Journal of Knowledge Management**, vol. 15, n°. 6, 2011.

REID, S. E.; BRENTANI, U. The fuzzy front end of new product development for discontinuous innovation: a theoretical model. **Journal of Product Innovation Management**, Malden, v. 21, n. 3, p. 170-184, 2004.

REPKO, A. F. **Interdisciplinary research**. Thousand Oaks: Sage Publications. 2008.

SABOL, V.; KIENREICH, W.; MUHR, M.; KLIEBER, W.; GRANITZER, M.. Visual Knowledge Discovery in Dynamic Enterprise Text Repositories. **13th International Conference Information Visualisation**, vol., n., p. 361-368, July 2009.

SANDSTRÖM C.; BJÖRK, J. Idea management systems for a changing innovation landscape. **International Journal of Product Development**, v.11, p. 310–324, 2010.

SCHREIBER, G. Knowledge engineering. In V. Lifschitz, F. van Harmelen, B. Porter (Eds). *Handbook of Knowledge Representation*. Elsevier, 2007. p. 929-946.

SCHREIBER, G., *et al.* **Knowledge Engineering and Management: the CommonKADS Methodology**. MIT Press, Cambridge, Massachusetts, 2002.

SCHUMPETER, J.A. The theory of economic development: an inquiry into profits, capital, credit, interest and the business cycle; translated

from the German by Redvers Opie. **Harvard University Press**, Cambridge, MA. 1949.

SCULLEY, D. Web-scale K-means *clustering*. In: **International Conference on World Wide Web**, 19., 2010, Raleigh, North Carolina, USA. Proceedings... Raleigh, North Carolina, USA, 2010.

SÉRGIO, Marina Carradore; GONÇALVES, Alexandre Leopoldo; SOUZA, João Artur de. UM MODELO PARA AUXILIAR NA TOMADA DE DECISÃO NO DOMÍNIO DE GESTÃO DE IDEIAS . 12ª Conferência Internacional Sobre Sistemas de Informação e Gestão de Tecnologia: CONTECSI. São Paulo, maio 2015. Disponível em: <<http://www.contecsi.fea.usp.br/envio/index.php/contecsi/12CONTECSI/paper/view/2978>>. Acesso em: 15 jun. 2015.

SILVA, E. R. G.; ROVER, A. J. O Processo de descoberta do conhecimento como suporte à análise criminal: minerando dados da Segurança Pública de Santa Catarina. In: International Conference on Information Systems and Technology Management, 2011, São Paulo. **Anais da International Conference on Information Systems and Technology Management**. São Paulo: FEA, 2011. v. 8.

SILVA, Edna Lúcia da; MENEZES, Estera Muszkat. **Metodologia da pesquisa e elaboração de dissertação**. 4. ed. rev. Atual, Florianópolis: UFSC, 2005, 138p

SINT, R.; MARKUS, M.; SCHAERT, S.; KURZ, T. Ideator - a collaborative enterprise idea management tool powered by KiWi. **Fifth Workshop "Semantic Wikis. Linking Data and People"**. Hersonissos, Greece, 2010.

STARBUCKS. **Ideas in action**. 2016. Disponível em: <<http://blogs.starbucks.com/blogs/customer/default.aspx>>. Acesso em: 18 fev. 2016.

STARBUCKS. **My Starbucks Idea**. 2015. Disponível em: <<http://mystarbucksidea.force.com>>. Acesso em: 22 jun. 2015.

STUDER, R.; BENJAMINS, V. R.; FENSEL, D. Knowledge engineering: principles and methods. **Data & knowledge engineering**, v. 25, n. 1, p. 161-197, 1998.

SUÁREZ-FIGUEROA, M. C.; GARCÍA-CASTRO, R.; VILLAZÓN TERRAZAS, B.; GÓMEZ-PÉREZ, A. Essentials In Ontology Engineering: Methodologies, Languages, and Tools. In: **CIB conference W078-W012 - 2nd Workshop on eeBuildings Data Models**, 26-28 out. 2011, Sophia Antipolis, França, 2011.

TAN, A.-H. Text mining: The state of the art and the challenges. In: **Proceedings of the Pacific Asia Conference on Knowledge Discovery and Data Mining – PAKDD'99 Workshop on Knowledge Discovery from Advanced Databases**, Beijing, p. 65–70, 1999.

TROTT, P. **Gestão da inovação e desenvolvimento de novos produtos**. 4. ed. Porto Alegre: Bookman, 2012.

VANDENBOSCH, B.; SAATCIOGLU, A.; FAY, S. Idea management: a systemic view. **Journal of Management Studies**, Malden, v. 43, n. 2, p. 259-288, 2006.

VANGUNDY, A. B. **Getting to innovation**: How asking the right questions generates the great ideas your company needs. New York: AMACOM, 2007.

VARSHOVI, A.; SADEGHIYAN, B. Ontological classification of network denial of service attacks: basis for a unified detection framework. **Scientia Iranica**, v. 17, n. 2 D, p. 133-148, 2010.

VASCONCELOS, E. S.; CRUZ, C. D., BHERING, L. L.; RESENDE JÚNIOR, M. F. R. Método alternativo para análise de agrupamento. **Pesq. agropec. bras.**, Brasília, v. 42, nº 10, 2007.

VASCONCELOS, José Braga de; ROCHA, Álvaro; KIMBLE, Chris. Sistema de informação de memória organizacional: uma abordagem ontológica para a definição de competências de grupo. In: **Conferência da Associação Portuguesa de Sistemas de Informação**, 4., 2003, Porto, Portugal. Actas... Porto, Portugal, 2003.

VASHISHTHA, J.; KUMAR, D.; RATNOO S. Revisiting Interestingness Measures for Knowledge Discovery in Databases. **Second International Conference on Advanced Computing & Communication Technologies**, p. 72–78. doi:10.1109/ACCT.2012.97, 2012.

VYGOTSKY, L. S. **A formação social da mente**. 7 ed. São Paulo: Martins Fontes, 2007.

WEISS, S. M., *et al.* **Text Mining: Predictive Methods for Analyzing Unstructured Information**. Spring, New York, 2005.

WESTERSKI, A.; DALAMAGAS, T.; IGLESIAS, C. A.. Classifying and comparing community innovation in Idea Management Systems, **Decision Support Systems**, 2013.

WESTERSKI, A.; IGLESIAS C. A.; RICO, F. T. Linked opinions: Describing sentiments on the structured web of data. *In 4th international workshop Social Data on the Web (SDoW2011)*, Bonn, Germany, October 2011.

WESTERSKI, A.; IGLESIAS, C. A.; RICO, F. T. A Model for Integration and Interlinking of Idea Management Systems. **4th Metadata and Semantics Research Conference (MTRS 2010)**, Alcalá de Henares, Spain, 2010.

WESTERSKI, Adam; IGLESIAS, Carlos A.; GARCIA, Javier Espinosa. Idea relationship analysis in open innovation crowdsourcing systems. In: **Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom), 2012 8th International Conference on**. IEEE, 2012. p. 289-296.

WESTERSKI, Adam; IGLESIAS, Carlos Angel. Exploiting Structured Linked Data in Enterprise Knowledge Management Systems: An Idea Management Case Study. EDOCW. p. 395-403, **IEEE Computer Society**, 2011.

WHITNEY, Dwight E. Assemble a technology development toolkit. **Research-Technology Management**, v. 50, n. 5, p. 52-58, 2007.

WITELL, Lars; LÖFGREN, Martin; GUSTAFSSON, Anders. Identifying ideas of attractive quality in the innovation process, **The TQM Journal**, Vol. 23, n. 1, p. 87-99, 2011.

WIVES, L. K. **Utilizando conceitos como descritores de textos para o processo de identificação de conglomerados (*clustering*) de documentos** – Tese (doutorado) – Universidade Federal do Rio Grande

do Sul. Programa de Pós-graduação em Computação, Porto Alegre, BR – RS, Brasil, 2004.

WROBLEWSKI, MICHAL. **A hierarchical www pages clustering algorithm based on the vector space model**. Tese de Doutorado. Poznań University of Technology, Poland. 2003.

XIE, L., ZHANG, P. Idea Management System for Team Creation. **Journal of Software**, North America, 5, nov. 2010.

ZANNI-MERK, C.; CAVALLUCCI, D.; ROUSSELOT, F. An ontological basis for computer aided innovation. **Computers in Industry**, v. 60, n. 8, p. 563-574, 2009.