

Willian Rochadel

**IDENTIFICAÇÃO DE CRITÉRIOS PARA AVALIAÇÃO DE
IDEIAS: UM MÉTODO UTILIZANDO FOLKSONOMIAS**

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

Orientador: Prof. Dr. João Artur de Souza

Coorientadora: Prof.^a Dr.^a Gertrudes Aparecida Dandolini

Florianópolis
2016

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

Rochadel, Willian

Identificação de critérios para avaliação de ideias : um método utilizando folksonomias / Willian Rochadel ; orientador, João Artur de Souza ; coorientadora, Gertrudes Aparecida Dandolini. - Florianópolis, SC, 2016.
177 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. Front End da Inovação. 4. Sistema de Conhecimento Coletivo. 5. Folksonomia. I. Souza, João Artur de. II. Dandolini, Gertrudes Aparecida. III. Universidade Federal de Santa Catarina. Programa de Pós Graduação em Engenharia e Gestão do Conhecimento. IV. Título.

Willian Rochadel

IDENTIFICAÇÃO DE CRITÉRIOS PARA AVALIAÇÃO DE IDEIAS: UM MÉTODO UTILIZANDO FOLKSONOMIAS

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

Florianópolis, 29 de fevereiro de 2016.

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

Banca Examinadora:

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

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

Prof. João Bosco da Mota Alves, Dr.
Universidade Federal de Santa Catarina

Prof. Roberto Tadeu Raittz, Dr.
Universidade Federal do Paraná

Este trabalho é dedicado aos meus queridos pais, Oliveira e Ruth.

AGRADECIMENTOS

Aos queridos orientadores, Prof. João Artur de Souza e Prof.^a Gertrudes Dandolini, pessoas sensacionais, ao qual me proporcionaram essa oportunidade, conduziram e orientaram nessa pesquisa, mas também me ensinaram muito com seu exemplo de dedicação, profissionalismo e carisma.

Ao Prof. Juarez Bento da Silva, que, desde a graduação, apoiou e me inseriu no meio científico por meio das pesquisas e desenvolvimentos no Laboratório de Experimentação Remota (RExLab), e aos colegas do grupo.

Aos membros da banca, pela participação, avaliação, considerações e recomendações para a pesquisa.

Aos demais docentes do Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento (PPGEGC), pelo alto nível de ensino e seriedade científica.

Aos colegas do PPGEGC e do Grupo de Pesquisa em Inteligência, Gestão e Tecnologia para a Inovação (IGTI), pela colaboração, acomodação e por tornar esse período um momento tão agradável.

E, em razão de surgir cada uma dessas pessoas ao meu lado, agradeço a Deus, que me sustentou até aqui.

*The most dangerous phrase in the language is,
"We have always done it this way".*

(Grace Hopper, 1987)

RESUMO

As ferramentas de cocriação encontram uma rica fonte de conhecimento baseada nas interações sociais que ocorrem na Web. Essa interação coletiva é a principal característica dos Sistemas de apoio à inovação, em especial para os sistemas de gestão de ideias. Entretanto, para avaliar ideias, as soluções atuais limitam-se a métodos baseados em formulários com critérios pré-estabelecidos ou, então, por métricas de engajamento social. O contexto organizacional é crítico para o sucesso de uma ideia, porém, ao considerar apenas índices de popularidade, as avaliações não agregam semanticamente o conhecimento atribuído pelo usuário, bem como não determinam quais critérios foram ponderados pela comunidade. A fim de compreender este conhecimento coletivo, a presente pesquisa propõe um método de identificação e análise de critérios para a avaliação de ideias. O desenvolvimento desse artefato é baseado na metodologia da ciência do design e explora o conhecimento a partir de atribuições sociais por notas e *tags*, as folksonomias. Assim, no contexto do *front end* da Inovação, o método representa uma apropriação semântica e qualitativa dos critérios atribuídos pela comunidade. A verificação utiliza técnicas da mineração de folksonomias em uma base de dados representada por um modelo de hipergrafo. Como resultado, o método permite evidenciar um conjunto de características a serem consideradas pela organização como critérios de avaliação. Além disso, a solução constata que a popularidade não é uma medida de consenso da comunidade, portanto sub comunidades auferem medidas mais precisas em suas atribuições; e a flexibilização temporal, própria das interações sociais, colaboram na recomendação de ideias baseada em tendências e no contexto organizacional.

Palavras-chave: Gestão de Ideias. *Front End* da Inovação. Sistema de Conhecimento Coletivo. Folksonomia

ABSTRACT

Co-creation tools meet a rich source of knowledge on social interactions that occurs on the Web. This collective interaction is the main characteristic of innovation support systems, especially idea management systems. However, in order to evaluate ideas, current solutions are limited to methods based on forms with pre-established criteria or metrics of social engagement. The organizational context is critical to the success of an idea. Nevertheless, when considering just popularity ratings, the evaluations do not semantically aggregate the knowledge attributed by the user. It also does not determine what criteria was weighted by the community. In order to understand this collective knowledge, the present research proposes a method for identification and analysis of criteria in idea evaluation. The development of this artefact is based on the design science research methodology, and it explores the knowledge from social attributions using grades and tags, also known as folksonomy. Therefore, within the front end of innovation, the method represents a semantic, qualitative appropriation of criteria attributed by the community. The artefact was verified using folksonomy mining techniques in a database represented by a hypergraph model. As a result, the method allows to visualize a set of characteristics to be considered as evaluation criteria by any organization. In addition, the results showed that popularity is not a community's consensus measure. Therefore, sub communities get more precise measurements in their attributes; and temporal flexibility, which is specific to social interactions, collaborate on the idea recommendation based on trends and organizational context.

Keywords: Ideas Management. Front End of Innovation. Collective Knowledge System. Folksonomy

LISTA DE FIGURAS

Figura 1 – O Paradigma de inovação aberta.	34
Figura 2 – Modelo de Desenvolvimento de novos conceitos.....	36
Figura 3 – Modelo " <i>House of innovation</i> ".	38
Figura 4 – Arquitetura de um IMS.....	42
Figura 5 – Dependências do ciclo de vida da ideia.	43
Figura 6 – Ciclo de vida da ideia.	44
Figura 7 – A seleção de ideias em um Sistema de Gestão de Ideias.	51
Figura 8 – Conceito de cenário planejado.....	52
Figura 9 – Escopo de compartilhamento, Estabilidade e Formalidade da informação.....	55
Figura 10 – Exemplo de um Sistema de Conhecimento Coletivo.....	58
Figura 11 – Ontologia Folksonomizada.....	60
Figura 12 – Modelo de grafo tripartido.	61
Figura 13 – Representação formal de marcações e folksonomias.	62
Figura 14 – Decomposição completa de um hipergrafo.	64
Figura 15 – Visão Geral das abordagens de recomendação de <i>tag</i>	70
Figura 16 – Modelo do processo em DSRM.....	78
Figura 17 – Visão geral do Método.	87
Figura 18 – Modelo para a representação do conhecimento.	90
Figura 19 – Atribuições de nota e etiquetas desassociadas, conforme conjunto de dados.....	98
Figura 20 – Atribuições de avaliação e etiquetas associadas, conforme o modelo proposto.....	99
Figura 21 – Junção das propriedades de diferentes arquivos na classe "Atribuição".....	99
Figura 22 – Adequação dos dados.....	100
Figura 23 – Filtros aplicados para a decomposição do grafo em subgrafo.	103
Figura 24 – Subgrafo do usuário selecionado.	103
Figura 25 – Subgrafo AT do usuário (U) com as etiquetas (T) utilizadas nas atribuições (A).	104
Figura 26 – Rede de afiliação do usuário U17597.....	105
Figura 27 – Grafo das cinco <i>tag</i> com maior grau da rede.	106
Figura 28 – <i>Tags</i> selecionadas por grau de centralidade local e seus respectivos nós vizinhos.	109
Figura 29 – Subgrafo dos recursos destacados pelo PageRank entre 2006-2008.	111
Figura 30 – Nuvem de <i>Tags</i> a partir da busca sobre Sistemas de Gestão de Ideias (IMSS).	151
Figura 31 – Nuvem de <i>tag</i> a partir da pesquisa na área de Folksonomias e grafos.	163

LISTA DE QUADROS

Quadro 1 – Relação dos procedimentos à pesquisa	81
Quadro 2 – Fundamentação teórica.	84
Quadro 3 – Arquivo e propriedades do conjunto de dados.	97
Quadro 4 – Publicações relacionadas a dissertação.	127
Quadro 5 - Fluxo da Revisão.....	147
Quadro 6 – Documentos por área do conhecimento.	150
Quadro 7 – Palavras-chave por ocorrência.	150
Quadro 8 – Publicações por instituição.	152
Quadro 9 – Fontes das publicações.....	153
Quadro 10 – Mapeamento dos Artigos de IMS e Dados Bibliométricos.	155
Quadro 11 – Documentos por área do conhecimento.	162
Quadro 12 – Palavras-chave por ocorrência.	162
Quadro 13 – Publicações por instituição.	164
Quadro 14 – Fontes das publicações.....	166
Quadro 15 – Mapeamento dos Artigos de Folksonomias e Dados Bibliométricos com duas ou mais citações.....	167
Quadro 16 – Pesquisa sobre "ideias" no BTD do PPGEGC.	175
Quadro 17 – Pesquisa sobre "tomada de decisão" no BTD do PPGEGC.	176
Quadro 18 – Pesquisa sobre "análise de rede" no BTD do PPGEGC.....	176

LISTA DE GRÁFICOS

Gráfico 1 – Notas no subgrafo de Atribuições.....	102
Gráfico 2 – Número de publicações por autores, com no mínimo duas publicações.	151
Gráfico 3 – Publicações por país.	152
Gráfico 4 – Artigos por ano de publicação.	153
Gráfico 5 – Número de publicações dos principais autores.	164
Gráfico 6 – Publicações por país.	165
Gráfico 7 – Artigos por ano de publicação.	165

LISTA DE ABREVIATURAS E SIGLAS

FEI – *Front End* da Inovação

Gi2MO – *Generic Idea and Innovation Management Ontology*

IGTI – Núcleo de Pesquisa em Inovação, Gestão da Tecnologia da Informação

IMS – Idea Management System; Sistema de Gestão de Ideias

PPGEGC – Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento

SUMÁRIO

1	INTRODUÇÃO	25
1.1	CONTEXTUALIZAÇÃO E PROBLEMA DE PESQUISA	25
1.2	OBJETIVOS	26
1.2.1	Objetivo geral	26
1.2.2	Objetivos específicos	27
1.3	JUSTIFICATIVA	27
1.4	ADERÊNCIA AO PPGEGC/UFSC	28
1.5	PROCEDIMENTOS METODOLÓGICOS	29
1.6	DELIMITAÇÕES DA DISSERTAÇÃO	30
1.7	ESTRUTURA DA DISSERTAÇÃO	30
2	ESTADO DA ARTE	33
2.1	A INOVAÇÃO	33
2.1.1	O <i>Front End</i> da Inovação	35
2.1.2	A atividade de Gestão de Ideias	37
2.1.3	As etapas de avaliação e seleção de ideias no FEI	38
2.1.4	Sistemas de apoio à Inovação	40
2.2	SISTEMAS DE GESTÃO DE IDEIAS – IMS	41
2.2.1	O Ciclo de vida das Ideias nos IMS	42
2.2.2	Os Aspectos da Gestão do Conhecimento no ciclo	45
2.2.3	A representação do conhecimento em sistemas Web	46
2.3	AS ATIVIDADES DE AVALIAÇÃO E SELEÇÃO EM IMS	48
2.3.1	Métodos e técnicas para a avaliação e seleção	49
2.3.2	O tratamento de ideias rejeitadas	51
2.4	REDES SEMÂNTICAS	53
2.4.1	Semântica na Web	54
2.4.2	Folksonomia	56
2.4.3	Conhecimento coletivo na Web	57
2.4.4	Representação das Folksonomias em Hipergrafos	60
2.5	ANÁLISE DE FOLKSONOMIAS	62

2.5.1	Decomposição	63
2.5.2	Classificação	66
2.5.3	Distribuição	67
2.5.4	Recomendação	69
2.5.5	As ferramentas de Análise	71
2.6	CONSIDERAÇÕES DO CAPÍTULO	72
3	PROCEDIMENTOS METODOLÓGICOS	75
3.1	METODOLOGIA DE PESQUISA DA DESIGN SCIENCE	75
3.1.1	As atividades do modelo	77
3.1.2	Os pontos de entrada do modelo	78
3.1.3	Os artefatos resultantes	79
3.2	RELAÇÃO DOS PROCEDIMENTOS À PESQUISA	81
4	APRESENTAÇÃO E ANÁLISE DOS ARTEFATOS	83
4.1	A ABORDAGEM DO CONHECIMENTO	83
4.2	A INTEGRAÇÃO DO MÉTODO AOS IMS	85
4.3	A CARACTERIZAÇÃO DO MÉTODO	86
4.3.1	Avaliação dos Usuários	88
4.3.2	Base do conhecimento	89
4.3.3	Análise das atribuições	92
4.3.4	Definições pela Organização	94
4.4	A VERIFICAÇÃO DO MÉTODO PROPOSTO	96
4.4.1	O Conjunto de Dados	96
4.4.2	A análise das folksonomias	101
4.4.3	Os Resultados	119
4.5	CONSIDERAÇÕES DO CAPÍTULO	123
5	CONSIDERAÇÕES FINAIS	125
5.1	REVISÃO DO PROBLEMA E DA PROPOSTA	125
5.2	GENERALIZAÇÃO	126
5.3	CONTRIBUIÇÕES	127
5.4	LIMITAÇÕES	128

5.5	RECOMENDAÇÕES E TRABALHOS FUTUROS	128
	REFERÊNCIAS	129
	APÊNDICE A – Metodologia de Revisão.....	147
	APÊNDICE B – Bibliometria de IMS	149
	APÊNDICE C – Bibliometria de Folksonomias	161
	APÊNDICE D – Trabalhos Relacionados ao EGC	175
	APÊNDICE E – Esquema Metodologia	177

1 INTRODUÇÃO

Este capítulo apresenta o tema do trabalho, a justificativa, os objetivos, as delimitações, a aderência ao Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento (PPGEGC) e a estrutura do documento.

1.1 CONTEXTUALIZAÇÃO E PROBLEMA DE PESQUISA

A Inovação é definida como um processo de várias etapas. Através destas etapas “*as organizações transformam ideias em produtos novos/melhorados, serviços ou processos, a fim de avançar, competir e diferenciar-se com sucesso em seu mercado*” (Baregheh *et al.*, 2009). Koen *et al.* (2014a) dividem este processo em três principais fases: o *front end* da inovação (FEI), o processo de desenvolvimento de novos produtos, e a comercialização ou implementação. Sendo que, por suas características de incerteza e riscos envolvidos, o *front end* é considerada a fase mais crítica (Koen *et al.*, 2014b).

O impacto dessas características se confronta com a importância do *front end* no processo estratégico da Inovação (Russell e Tippett, 2008). Teza (2012) identifica que os estudos desta fase estão mais voltados a atividade de “*geração de ideias, com foco em fontes de ideias e/ou nas técnicas de geração de ideias*”, e identifica a carência de estudos nas demais atividades.

No processo criativo de geração de ideias, um grande número de alternativas pode ser considerado. Alternativas descritas apenas com um título e uma curta caracterização, na maioria das vezes. Portanto, um processo criativo pode gerar o máximo de possibilidades/ideias possíveis.

Kempe *et al.* (2012) identificam estas alternativas com o termo em inglês “*raw idea*” para expressar uma ideia em seu estado bruto, ou seja, ainda “*crua*” e descrita rapidamente. Herstatt *et al.* (2006) definem esta carência de informação como um “*fator limitante*” para o processo de Inovação como um todo e principalmente para o FEI. Isso pode acarretar em duas grandes dificuldades: (a) uma lacuna no processo de seleção; e/ou (b) incerteza no resultado da seleção geral.

Kempe *et al.* (2012) aborda o FEI em duas fases principais. Em uma primeira fase divergente/generalista, as ideias são geradas e/ou recolhidas. Posteriormente, em uma fase convergente estas ideias são enriquecidas, avaliadas e selecionadas iterativamente e interativamente.

Nessa convergência os fatores limitam o julgamento a partir de critérios objetivos, e as impressões subjetivas dos avaliadores são obtidas. Assim, a alternativa das decisões é baseada no critério “*X é melhor do que Y*” na comparação entre pares (Horton e Goers, 2014).

Estas impressões subjetivas geram um dos problemas centrais no FEI: as incertezas (Schmitt, 2013). Andrade (2011) analisa que “*o conceito de incerteza está associado à ausência (ou limitação) de conhecimento e é contrastado com a ideia de conhecimento na forma de cálculo de risco probabilístico quantificável*”.

Considerar o conhecimento e incertezas é um desafio para a tomada de decisões. Pois, nesta fase crítica da inovação erros podem resultar em uma ideia de possível valor rejeitada sem retorno, ou fazer com que uma ideia consuma tempo e recurso resultando em um fracasso (Brun *et al.*, 2009). O que corresponde aos riscos envolvidos no FEI que o modelo *Stage Gate*[®] (Cooper, 2008) associa ao custo crescente de pesquisa e desenvolvimento em cada etapa.

Portanto, em relação a ideias e propostas, as incertezas estão diretamente relacionadas com os critérios de avaliação tomados e, a partir disso, quais delas realmente estão hábeis para serem selecionadas, ou seja, que valerão esforço, tempo e recursos empreendidos para serem enriquecidas, desenvolvidas ou pesquisadas.

Baseando nestes pontos levantados pela literatura, é definida a pergunta de pesquisa: “**Como identificar e analisar critérios para avaliação de ideias a partir do conhecimento coletivo em um sistema de gestão de ideias?**”.

1.2 OBJETIVOS

Considerando a problemática discutida, são descritos os objetivos da pesquisa a seguir.

1.2.1 Objetivo geral

Propor um método de identificação e análise de critérios para avaliação de ideias a partir do conhecimento coletivo em um sistema de gestão de ideias.

1.2.2 Objetivos específicos

- a) Identificar critérios de análise de ideias;
- b) Analisar o conhecimento coletivo na avaliação de ideias;
- c) Avaliar ideias em diferentes contextos;

1.3 JUSTIFICATIVA

Este estudo faz parte das pesquisas do IGTI (Grupo de Pesquisa em Inteligência, Gestão e Tecnologia para a Inovação) em relação ao *front end* da Inovação, entre os temas a Gestão de Ideias, considerando os trabalhos de Prada (2009), Miguez (2012) e Dorow (2013).

Uma inovação é iniciada por uma ideia (Mikelsons e Lielā, 2015). Uma ideia pode estar numa frase, num parágrafo ou ser um rabisco qualquer, porém uma simples ideia pode conceber um valor imprescindível para a organização (Kempe *et al.*, 2012). Nesse contexto, o conhecimento é a peça fundamental para avaliação e tomada de decisão em uma perspectiva global e específica por área, reconhecendo suas deficiências, necessidades e valores (Teza, 2012). A etapa de avaliação de ideias envolve características de avaliação de risco em diversas áreas do conhecimento, nestes casos a avaliação por especialistas das áreas específicas permite uma ampliação de aportes para tomada de decisão.

O importante desta etapa inicial é selecionar as ideias certas, executá-las de maneira coerente e com qualidade para alcançar os objetivos, e, desta forma, trazer resultados para a organização (Prada, 2009). Portanto o conhecimento de grupos de especialistas contribui em diferentes perspectivas para uma avaliação detalhada.

Em uma concepção geral, a tomada de decisão considera o contexto organizacional, enquanto as áreas operacionais e técnicas consideram outras perspectivas como as competências, maquinários e recursos disponíveis, já a área de marketing e mercado percebe as dimensões relacionadas aos custos, lucros e economia. Portanto, cada área possui uma especificidade, sendo fundamental uma avaliação que considere o conhecimento dos envolvidos em cada dimensão.

Os diversos modelos de gestão do portfólio apresentam a gestão de ideias como fator decisivo para início do processo de Inovação. Porém estes modelos carecem de detalhes e relacionamento com o planejamento estratégico e a identificação de oportunidades (Prada, 2009). O processo de gestão de ideias sem o alinhamento com os processos estratégicos fica limitado a criação. Por outro lado, a

utilização do conhecimento na tomada de decisão é um importante subsídio para auxiliar as definições estratégicas e gerar novas ideias. Assim, a ligação entre ideias e oportunidades com a dimensão estratégica possuem forte relacionamento para o auxílio a tomada de decisões.

Porém, os modelos comuns apresentam uma visão matricial de avaliação da ideia e, portanto, a avaliação do consultor se torna um fim em si mesmo, deixando de lado o valor da experiência, conhecimento e competência necessários para cada um avaliar cada critério.

1.4 ADERÊNCIA AO PPGE/C/UFSC

Como principal objeto de pesquisa tratado no Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, o conhecimento é identificado nesta pesquisa como um recurso essencial, porém complexo. Para isto são tratados os métodos e técnicas da Engenharia e Gestão do conhecimento em uma modelagem de natureza interdisciplinar baseada em Schreiber (2000).

Ao considerar que o conhecimento “*é conteúdo ou processo efetivado por agentes humanos ou artificiais em atividades de geração de valor científico, tecnológico, econômico, social ou cultural*”, Pacheco (2014) se refere a: (a) ao **propósito**, como a geração de valor; (b) a **forma**, como o conteúdo ou processo; (c) o **local**, considerando tanto a mente humana quanto em artefatos; (d) a **fonte**, como indivíduos, grupos, organizações, redes, regiões e países num contexto de inovação; e, por fim, (e) a **epistemologia**, que converge na convivência interdisciplinar.

Os estudos de Repko (2011) consideram a interdisciplinaridade como “*um processo de responder a uma questão, resolução de um problema ou a abordagem de um tema que é muito amplo ou complexo para ser tratado de forma adequada por uma única disciplina ou profissão*”. Considerando então que “*se baseia em perspectivas disciplinares e integra os seus conhecimentos para produzir uma perspectiva mais abrangente*”. Diferenciando-se da multidisciplinaridade devido aos “*mecanismos do processo de pesquisa e do produto final*”.

Ainda segundo Repko (2011), a interdisciplinaridade instrumental é orientada a problemas. Portanto, uma abordagem pragmática concentrada em pesquisa, suprimento, e na resolução de problemas práticos como resposta às demandas externas da sociedade. No entanto,

tomada sozinha, não é suficiente para a interdisciplinaridade instrumental, mas deve ser complementada pela integração das áreas, portanto *"to make whole"*, ou seja, "fazer o todo". No contexto de interdisciplinaridade, a integração é o processo pelo qual ideias, dados e informações, métodos, ferramentas, conceitos, e ou teorias de duas ou mais disciplinas são sintetizadas, conectadas, ou misturadas (Repko, 2011).

A visão integradora do contexto interdisciplinar está de acordo com o entendimento da Capes (2009) que define a área interdisciplinar *"entre o teórico e o prático, entre o filosófico e o científico, entre ciência e tecnologia, apresentando-se, assim, como um saber que responde aos desafios do saber complexo"*. Segundo Pacheco (2014), a proposta de definição interdisciplinar do conhecimento no escopo do programa EGC, permite a integração convergente das epistemologias do conhecimento, sendo definidos como: (a) **cognitivista**, *"o conhecimento pode ser conteúdo e estar em artefatos"*; (b) **conexionista**, *"rede e comunicação são essenciais na geração de valor"*; e (c) **autopoietico**, *"a criação de conhecimento se dá, principalmente, por humanos"*.

Sendo apontadas como disciplinas bases: a gestão do conhecimento, na análise dos domínios e critérios de avaliação; e a engenharia do conhecimento, na análise e representação do conhecimento. Logo se integra ao programa na concepção de: *"Aplicação de sistemas de conhecimento à gestão do conhecimento organizacional (formalização, memória e tomada de decisão);"* (EGC, 2004) da linha de pesquisa "Engenharia do Conhecimento aplicada às organizações".

Por fim, cabe ressaltar a continuidade de pesquisas que já vem sendo realizadas no âmbito do EGC e principalmente do grupo de pesquisa IGTI, vinculado ao EGC, como: Prada (2009), Teza (2012), Miguez (2012), Dorow (2013), Fernandes (2012), Schmitt (2013), Wilbert (2015) e Buchele (2015).

1.5 PROCEDIMENTOS METODOLÓGICOS

Devido à articulação interdisciplinar do processo teórico-metodológica que proporciona a viabilidade da abordagem ao problema de pesquisa e situa a compreensão de uma adoção pragmática.

A Visão de mundo deste projeto é o Paradigma Funcionalista (Burrell e Morgan, 1979) pela orientação pragmática ao qual a compreensão do estudo deve ser posta em termos de conhecimentos

gerais e posteriormente em prática, preocupando-se em uma solução funcional para o problema, com teorias já existentes.

Esta pesquisa representa a modalidade tecnológica por seu objetivo em avançar a tecnologia na solução. Para isto explora uma abordagem metodológica qualitativa a partir de uma revisão integrativa atualizada sobre os métodos e ferramentas que poderão ser integrados ao modelo tecnológico.

Sobre estes aspectos, segue a metodologia do *Design Science Research* (Peppers *et al.*, 2007) para o levantamento do estado da arte e definição do modelo.

1.6 DELIMITAÇÕES DA DISSERTAÇÃO

Esta pesquisa tem como tema a etapa inicial do processo de Inovação, definido na literatura como *front end* da Inovação (FEI). Dentro das pesquisas do FEI se limita a atividade de Gestão de Ideias, apesar de identificar que essa atividade se estende por todo o processo.

Nesse tema, como parte dos estudos mais amplos do grupo IGTI, a abrangência se delimita pelo objetivo da dissertação em abordar a avaliação de ideais. Outras questões relacionadas ao armazenamento de informações e tratamento de tomadas de decisão não fazem parte do escopo apresentado. E, portanto, são questões fundamentais no tema, mas apontadas como estudos futuros.

1.7 ESTRUTURA DA DISSERTAÇÃO

A presente dissertação está estruturada em cinco capítulos sendo:

- a) Esse primeiro capítulo introduz o tema da dissertação, constata o problema de pesquisa e descreve o objetivo geral e os específicos, a justificativa, a aderência do tema ao PPGEGC, bem como as delimitações e a estrutura do trabalho;
- b) O segundo capítulo apresenta uma revisão sobre os temas de Inovação, Sistemas de Gestão de Ideias e Redes Semânticas, apresentando os conceitos que permitem o embasamento teórico do modelo;
- c) No terceiro capítulo são descritos os procedimentos metodológicos utilizados para o desenvolvimento dos artefatos.

- d) O quarto capítulo então descreve os artefatos, as análises e as principais constatações;
- e) Por fim, o quinto capítulo conclui a dissertação apresentando as considerações finais sobre o tema, contribuições, limitações e as recomendações para trabalhos futuros.

2 ESTADO DA ARTE

Neste capítulo são apresentados os conceitos centrais para a fundamentação da pesquisa considerando o conhecimento e a interdisciplinaridades das áreas de Inovação, Sistemas de Gestão de Ideias; Redes e Folksonomias.

2.1 A INOVAÇÃO

A inovação é um tema central cada vez mais discutido pelas organizações e pesquisado pela academia. Isso devido a sua relevância em um mercado dinâmico e altamente competitivo.

Existem muitas definições de inovação, tanto do ponto de vista de resultado como de processo. O Manual De Oslo (2004) define a inovação como o resultado, seja em produto, processo, marketing ou método organizacional, como citado na introdução. Os estudos de Vandebosch *et al.* (2006) apresentam a inovação como “*a implementação de uma ideia criativa que pode ser expressa na forma de conhecimento, de uma prática ou de um objeto físico*”. Já Baregheh *et al.* (2009) define a inovação como um processo: “*a inovação é um processo de várias etapas através do qual as organizações transformam ideias em produtos novos/melhorados, serviços ou processos, a fim de avançar, competir e diferenciar-se com sucesso em seu mercado*”.

Logo, é observado nessas definições que existe uma complementaridade nos conceitos, afinal, o processo deve gerar um resultado, uma inovação. Entretanto, a palavra resultado pode ser interpretada de várias formas e opta-se nesta pesquisa pela visão da inovação como um processo de várias etapas que precisa ser gerenciado para minimizar riscos e aumentar as possibilidades de sucesso expresso por um conhecimento, prática ou objeto novo.

Portanto, o processo de Inovação para ser eficiente deve ser precedido pela aquisição de um conhecimento novo. Porém, esse processo é custoso, dispende tempo e apresenta riscos (Chibás *et al.*, 2013). De acordo com Bessant e Tidd (2009), manter equipes de pesquisa e desenvolvimento também é dificultoso para empresas, tanto as de porte menor quanto grandes empresas.

Assim, segundo Chesbrough (2003), as competências internas da organização já não são mais suficientes, sendo necessário considerar uma abertura para que as inovações fluam entre o meio interno e externo.

Chesbrough (2003) cunhou o termo “Inovação aberta”, ou “*Open Innovation*”, que é definido como um paradigma sobre a permeabilidade

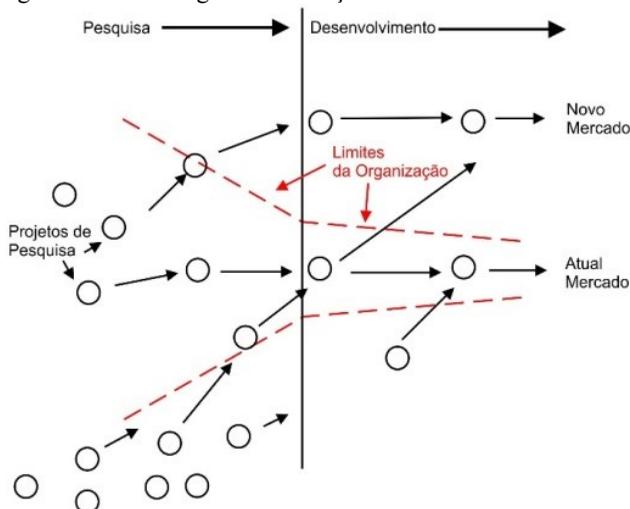
dos limites organizacionais movido por um sistema relacional que compreende a organização e seus parceiros externos, contrários ao modelo fechado.

Lindergaard (2011) diferencia os paradigmas de Inovação Fechada e Inovação Aberta pelo modo que é realizada a atividade de seleção de novas ideias. Enquanto a primeira tem por foco o ambiente interno, a Inovação Aberta realiza uma integração entre ideias e tecnologias externas a organização. O principal propósito disso é fomentar um potencial em novas áreas de atuação ou em ambientes com novos conhecimentos a serem explorados pela organização. Nesse contexto, a ideia ou as ideias passam a ser o centro e a organização apenas um dos agentes.

Chesbrough (2003) apresenta o processo de inovação aberta de três modos, conforme Figura 1:

- De fora para dentro** – processo de entrada que envolve a entrada e aquisição de conhecimento a partir de fontes externas à organização;
- De dentro para fora** – processo de saída, o qual envolve a saída e comercialização do conhecimento;
- Ambos** – processo de junção, que combina os processos de entrada e saída para resultar em uma cocriação de conhecimento.

Figura 1 – O Paradigma de Inovação Aberta.



Fonte: traduzido de Chesbrough (2003).

No paradigma da Inovação Aberta, dentre os desafios identificados se destacam: a necessidade da melhoria dos fluxos de conhecimento (Taglino *et al.*, 2012); o compartilhamento do conhecimento dentro e fora das organizações (Carbone *et al.*, 2010); o aumento nas atividades de pesquisa; e a relevância da formação de conexões e redes (Chen *et al.*, 2011; Zhang *et al.*, 2011; Poveda *et al.*, 2012).

2.1.1 O *Front End* da Inovação

O processo de inovação pode ser entendido por etapas a serem vencidas. A fase inicial, chamada de *front end*, é abordada como a mais difícil e, portanto, abordada mais detalhadamente para essa pesquisa. Nesta fase é que são feitas as proposições de ideias ou soluções para determinado mercado, ou necessidade específica dos clientes ou mesmo da própria organização.

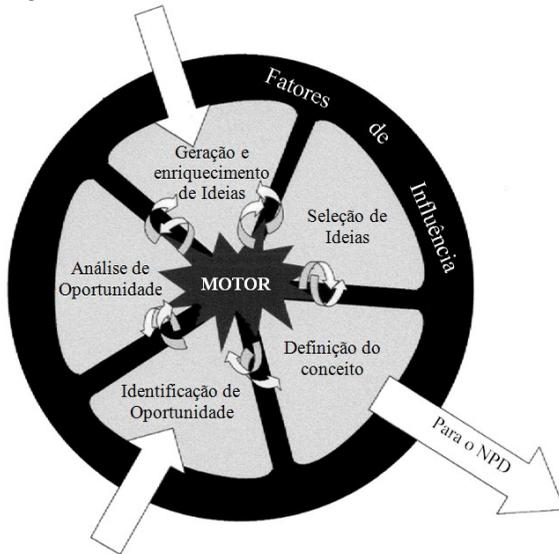
Nessa direção, Koen *et al.* (2002) dividem o processo da inovação em três principais fases: (a) o *fuzzy front end* da inovação; (b) o processo de desenvolvimento de novos produtos; e (c) a comercialização ou implementação. Porém, o termo “Fuzzy” transmite uma concepção de que esta fase não pode ser gerenciada, devido a fatores incontrolláveis. Esta fase inicial do processo de inovação realmente é caracterizada pela incerteza, porém é gerenciável e pode ser formalizada (Flint, 2002). Além de *Fuzzy Front End*, Teza (2012) ainda descreve outros termos utilizados para se referir a essa fase: *Front End Process*, *Front End of New Product Development*, *Early Phases of Innovation* e *Pre-development*.

Tendo em vista que a gestão da inovação nessa fase é determinante para o sucesso do processo como um todo, Koen *et al.* (2001) utilizam apenas o termo “*front end* da inovação” (FEI) defendendo uma gestão estratégica e complexa voltada aos objetivos e propósitos da organização.

O *front end* da inovação (FEI) é a fase em que uma oportunidade é considerada pela primeira vez e quando uma ideia é julgada pronta para o desenvolvimento (Kim e Wilemon, 2002).

Para a fundamentação da presente pesquisa será adotado o modelo de desenvolvimento de novos conceitos (*New Concept Development Model* – NCD) definido por Koen *et al.* (2002), conforme Figura 2.

Figura 2 – Modelo de Desenvolvimento de novos conceitos.



Fonte: Traduzido de Koen et al. (2002).

O modelo NCD, apresentado por Koen et al. (2002), expõe uma forma não sequencial, mas interativa entre os elementos, as setas indicam este movimento. As entradas podem ser ideias ou oportunidades, e a saída, um novo conceito para o desenvolvimento de novos produtos (NCD). O modelo NCD é dividido em:

- **Motor:** parte que dirige as cinco atividades do FEI. É o motor que move e dinamiza as atividades do FEI, sendo formado pela liderança, cultura e estratégias de negócio, portanto, o suporte da alta gestão;
- **Fatores de influência:** relacionam todas as outras partes do modelo. Esta parte afeta as decisões e são compostas das capacidades organizacionais, estratégia de negócios, do mundo exterior (como fornecedores, distribuidores, clientes e concorrentes) e também os avanços da ciência e da tecnologia que serão utilizados;
- **Área interior:** composta por cinco atividades: identificação de oportunidades, análise de oportunidade, geração de ideias, seleção de ideias e definição do conceito e tecnologia.

Com destaque às atividades apresentadas na Área interior do modelo, a sistematização da gestão de ideias é de extrema relevância para os processos de inovação, pois é a fase que as ideias são geradas e apropriadas (Aznar, 2011).

2.1.2 A atividade de Gestão de Ideias

No contexto do FEI, Kempe *et al.* (2012) identificam uma ideia como uma proposta crua, normalmente com um título e uma pequena descrição em uma fase de conhecimento generalista. Porém essa é a fase inicial de um processo ainda mais complexo definido na literatura como Gestão de ideias. Equivalentes às atividades de “Geração e enriquecimento de Ideias” e “Seleção de Ideias” no modelo NCD de Koen *et al.* (2002).

A Gestão de Ideias é considerada por Sandström e Björk (2010) como o processo pela qual os indivíduos e as empresas geram ideias criativas e as desenvolve. Para Kock *et al.* (2015), a definição e formulação de uma estratégia de gestão de ideias são significativas para que as organizações alinhem explicitamente seus processos de geração e seleção de ideias com suas estratégias de inovação.

Os fatores de sucesso do processo contribuem para a aquisição de conhecimento e reforço de uma cultura organizacional de inovação aberta, o que facilita a integração de muitas das partes interessadas no processo de ideação (Riel *et al.*, 2013).

Segundo Koen *et al.* (2001), neste processo uma ideia pode prover uma oportunidade como também uma oportunidade pode gerar um momento para uma ideia. O incentivo à produção de ideias e uma gestão voltada à inovação fomentam o potencial de criação e destaque da organização (Björk *et al.*, 2010). O resultado destas ideias inovadoras pode ser percebido em todo o processo, normalmente na forma de um conhecimento conceitual.

Engel *et al.* (2010) representam isso em um modelo em forma de pirâmide em três níveis: I) Inovação Estratégica; II) Inovação organizacional e cultural; e III) Ciclo de vida da gestão da Inovação. Na base da pirâmide denominada “*House of Innovation*” a gestão do conhecimento aparece como fator habilitador para a inovação estratégica e a gestão de ideias como um dos pilares do primeiro nível (Figura 3).

Figura 3 – Modelo "House of innovation".



Fonte: Traduzido de Engel *et al.* (2010).

2.1.3 As etapas de avaliação e seleção de ideias no FEI

Ao considerar a inovação como um processo (Baregheh *et al.*, 2009), o *front end* como uma fase (Koen *et al.*, 2014a), a gestão de ideias como um subprocesso (Sandström e Björk, 2010), as atividades de avaliação e seleção são apresentadas como etapas distintas.

No processo de Gestão de Ideias, a geração de ideias é apresentada como relativamente simples se comparada com as demais etapas. Em pouco tempo despendido, grupos e indivíduos qualificados são capazes de pensar em ideias considerando determinada área. Porém, e em especial se não considerar a interdisciplinaridade na criação, parte da geração será trivial, impraticável para a produção ou desprovida de conhecimento para ser desenvolvida (Prada, 2009).

Portanto, as etapas de avaliação e seleção são delicadas no processo. Há incertezas e ambiguidades e qualquer erro pode desprezar uma grande ideia ou dar continuidade a outra irrelevante (Brun *et al.*, 2009). A fim de melhor compreender e avaliar cada uma, é importante diferenciar as etapas de avaliação e seleção de ideias.

A etapa da avaliação deve ser implementada no contexto de compreensão do processo e dos objetivos da organização (Kock *et al.*, 2015). Com este intuito, a maioria dos métodos e técnicas utilizados

considera critérios que dimensionam a importância, a viabilidade, o descarte ou a junção das ideias (Alves *et al.*, 2007).

O valor de uma ideia está relacionado à identificação de uma oportunidade (Prada, 2009), portanto, é necessário listar critérios claros na avaliação. Para explorar o potencial é conveniente analisar o momento adequado, o recurso necessário, os drives estratégicos e o tempo para um amadurecimento e então o desenvolvimento. Neste processo, é significativo frisar a importância de uma cultura organizacional voltada à inovação, porém que considere os riscos e potencialidades (Riel *et al.*, 2013).

Após a avaliação, a etapa a seguir é a seleção das ideias. A triagem que ocorre especialistas é definida por critérios específicos para não serem analisadas apenas por sentimentos, sensibilidade ou pela percepção do grupo de consultores (Ferioli *et al.*, 2010). Estes critérios seguem as oportunidades percebidas pela organização, como as políticas internas e as externas. Além de considerar o julgamento da avaliação, na etapa de seleção são relacionados os objetivos estratégicos como auxílio à tomada de decisão. Envolvem, assim, os aspectos relacionados à ideia e também a cultura e estratégia organizacionais (Engel *et al.*, 2010).

A intenção da tomada de decisões é seguramente eleger as melhores ideias para serem desenvolvidas no portfólio de projetos inovadores (Alves *et al.*, 2007). Reduzir as incertezas e disponibilizar ferramentas que explorem o conhecimento é o propósito dos estudos voltados às etapas de avaliação e seleção de ideias.

Ferioli *et al.* (2010) ainda constata algumas percepções no seu estudo demonstrando dois distintos estágios na avaliação de ideias: a primeira utiliza critérios estruturados, um filtro de viabilidade técnica; a segunda é baseada em critérios subjetivos e no “instinto”. O autor constata que, em geral, especialistas passam mais tempo com ideias que serão aceitas do que as que serão rejeitadas.

Os dados da pesquisa também mostram uma nítida diferença na mentalidade dos participantes, após uma hora de avaliação, muito mais ideias são rejeitadas e que, portanto, sessões de avaliação de ideias longas não são muito benéficas. Portanto, o auxílio a tomada de decisão deve considerar estes erros e apoiar medidas para reduzir os impactos na continuidade do processo de inovação.

Ferioli *et al.* (2010) apresentam um framework baseado em critérios objetivos e subjetivos. Os critérios objetivos são argumentos utilizados para fazer uma avaliação quando uma evidência de raciocínio existe, nesta categoria os especialistas procuram tomar decisões baseadas em fatos concretos e relevantes. Na categoria de critérios

subjetivos ocorre uma avaliação estimada, feita por um julgamento baseado em impressões ou opiniões. Esta categoria é ainda dividida em três partes: “critérios subjetivos” que fica a cargo do próprio avaliador; “avaliação de sentimento” que são as decisões tomadas sem critérios, que não podem ser explicadas pelo raciocínio e estão além da argumentação; e por último a “decisão instantânea” que ocorre quando uma decisão é tomada de forma rápida e sem discussão.

2.1.4 Sistemas de apoio à Inovação

O paradigma da inovação aberta ampliou a preocupação das organizações em melhor gerir o conhecimento produzido pelo processo. A fim de auxiliar nessa gestão, os sistemas de inovação aberta apresentam alternativas para as organizações aproveitarem o conhecimento externo nos processos de inovação.

Hrastinski *et al.* (2010) apresenta uma revisão que identifica quatro categorias de sistemas de apoio à inovação:

- **Sistema de Gestão de Ideias:** Permitem que os usuários sugiram, avaliem e discutam ideias abertamente ou dentro de categorias pré-definidas;
- **Sistemas de resolução de problemas:** Oferecem oportunidades para a definição de problemas e, em seguida, sugerindo, avaliar e debater soluções;
- **Marketplace da Inovação:** pede aos usuários para sugerir soluções para os problemas definidos por uma organização, e usar recompensas e reconhecimento como um incentivo;
- **Sistemas de análise da inovação:** Fornecem ferramentas sofisticadas para avaliar e analisar a qualidade e potencial de ideias e soluções.

A revisão de Hrastinski *et al.* (2010) identifica que a principal característica dos sistemas de inovação aberta é a maneira como o conteúdo é avaliado, sendo considerada a participação da comunidade.

Outros sistemas comuns adquirem formatos voltados ao processo de gestão da inovação como: fóruns (Bettoni *et al.*, 2010; Westerski e Iglesias, 2011; De Liddo *et al.*, 2012; Spencer, 2012); blogs (Golder e Huberman, 2006; Krause *et al.*, 2008; Westerski e Iglesias, 2011; De Liddo *et al.*, 2012); *microbloggings* (Penela *et al.*, 2011); *wikis* (Baloh *et al.*, 2007; Jäschke *et al.*, 2008; Bailey e Horvitz, 2010; Bettoni *et al.*,

2010; De Liddo *et al.*, 2012; Strohmaier *et al.*, 2012); redes sociais (De Liddo *et al.*, 2012); Moodle (*Modular Object-Oriented Dynamic Learning Environment*) (Bettoni *et al.*, 2010; Bettoni *et al.*, 2013); mensageiros instantâneos (De Liddo *et al.*, 2012); entre outros.

2.2 SISTEMAS DE GESTÃO DE IDEIAS – IMS

A redução de investimentos em meio à crise financeira global em 2007 impactou diretamente os processos de inovação (Kanerva e Hollanders, 2009). Segundo Westerski (2013), o fato é que a motivação para inovar mudou, voltando-se então do propósito de exclusivamente servir para o aumento da competitividade ou a satisfação dos clientes, para se transformar em uma ferramenta de redução dos custos de produção.

Desde então há um avanço na utilização de tecnologias digitais para a gestão de ideias no processo de inovação, em especial na Web. O sucesso desses sistemas que aproveitam a criação de ideias a partir das multidões, ou então o paradigma da inovação aberta, gerou um grande volume de dados, muitos dos quais concebidos de maneira não estruturada e assim complexos de gerenciar (Perez *et al.*, 2014). Consequentemente surge a necessidade de um sistema apropriado para o gerenciamento e acompanhamento das ideias.

Os Sistemas de Gestão de Ideias (em inglês *Idea Management System – IMS*) são plataformas de software ou aplicativos que fornecem repositórios e ferramentas para coletar, procurar, editar, comentar e votar ideias (Perez *et al.*, 2014). Acomodam diferentes concursos ou campanhas e *stakeholders* como operadores, especialistas, tomadores de decisão e administradores (Bettoni *et al.*, 2010). Geralmente abordam uma implementação do conceito de inovação aberta no ambiente Web com o uso de técnicas do *crowdsourcing* (Westerski *et al.*, 2012).

Assim, o objetivo de tais sistemas é “*organizar a entrada, avaliá-la e produzir uma lista das melhores ideias que irão potencialmente entregar benefícios para a organização*” (Westerski *et al.*, 2012).

Segundo Sandström e Björk (2010), os sistemas de gestão de ideias podem ser vistos como o suporte estruturado para o processo de ideação.

Os tópicos a seguir apresentam a análise do termo na literatura e discute as características e conceitos encontrados.

2.2.1 O Ciclo de vida das Ideias nos IMS

O ciclo de vida das ideias pode ser definido como um processo em etapas sequenciais de amadurecimento e evolução, partindo de um momento de criação, armazenamento, enriquecimento e uso (Kock *et al.*, 2015).

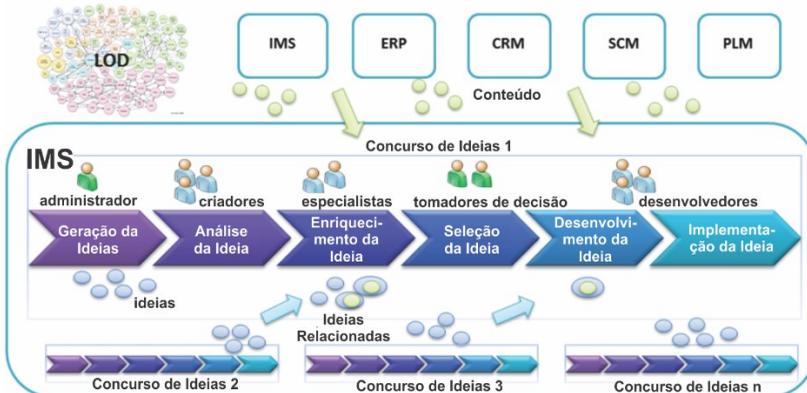
As diferentes abordagens do ciclo de vida da ideia em um IMS demonstram a importância de uma gestão eficaz dos dados gerados e do conhecimento que é desenvolvido nas fases. O que justifica a importância de um sistema apropriado para o contexto organizacional.

Segundo Perez *et al.* (2014) um IMS deve fornecer fluxos de trabalho para gerenciar as interações entre as diferentes fases do processo de inovação. O autor apresenta então um ciclo em seis etapas: geração de ideias, análise, enriquecimento, seleção, desenvolvimento e implementação (Figura 4).

Nessa arquitetura, Perez *et al.* (2014) reforça o aproveitamento das bases de dados dos sistemas corporativos como ERPs (*Enterprise Resource Planning*), CRMs (*Customer Relationship Management*), SCMs (*Supply Chain Management*), PLMs (*Product Lifecycle Management*) e outros IMSs, além da ligação de dados externos (LOD).

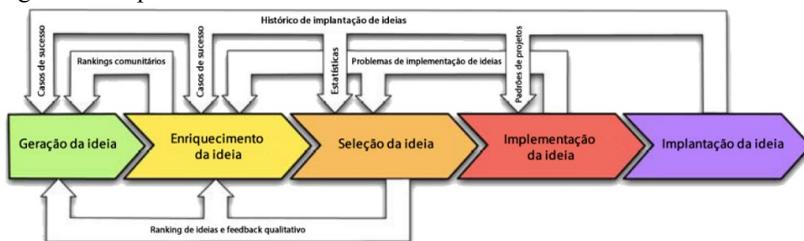
Na arquitetura apresentada por Perez *et al.* (2014), cada fase é suportada por atores com funções distintas: administrador, criadores, especialistas, tomadores de decisão e desenvolvedores. A arquitetura também define que ideias fluem sequencialmente em diferentes contextos e paralelamente entre elas no processo de enriquecimento.

Figura 4 – Arquitetura de um IMS.



Fonte: Traduzido de Perez *et al.* (2014).

Figura 5 – Dependências do ciclo de vida da ideia.



Fonte: Traduzido de Westerski *et al.* (2011).

O ciclo de ideias nos IMSs também é discutido por Westerski *et al.* (2011) que apresenta um fluxo iterativo e interativo em cinco fases. Dentre estas fases, o fluxo de informações fomenta o ciclo em cada etapa, mantendo o histórico de pesquisa e desenvolvimento, conforme pode ser observado na Figura 5.

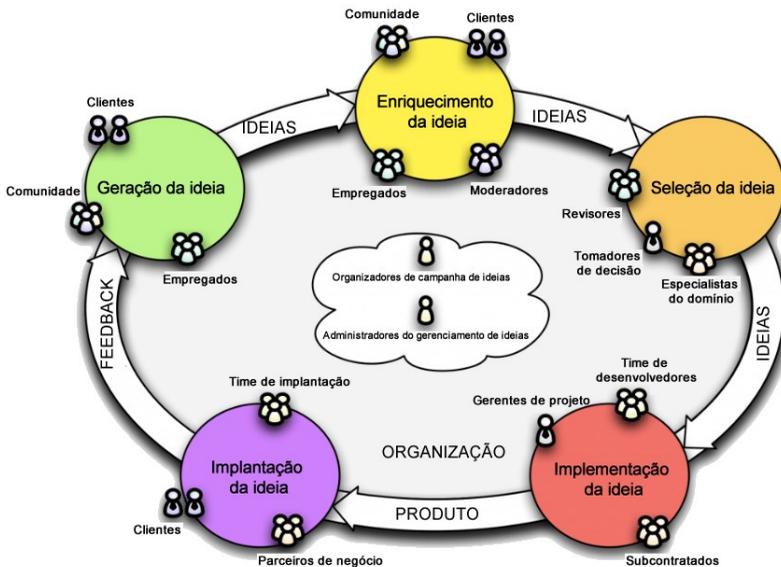
As cinco fases do ciclo de vida do IMS de Westerski *et al.* (2011) e Westerski *et al.* (2010) compreendem:

- a) **Geração da Ideia (“Idea generation”)**: extração de ideias de um grupo ou comunidade;
- b) **Aperfeiçoamento da Ideia (“Idea improvement”)**: colaboração entre os usuários para aperfeiçoar as ideias recolhidas. Gera um ranking pela comunidade utilizado como informação para a fase “a”;
- c) **Seleção da Ideia (“Idea selection”)**: aproveitamento dos dados gerados e escolha das melhores ideias. A interação gera o ranking das ideias e *feedbacks* para as fases “a” e “b”;
- d) **Aplicação da Ideia (“Idea implementation”)**: transformação de uma ideia em produto ou serviço. Os problemas que forem identificados retornam como informação para as etapas “b” e “c”;
- e) **Desenvolvimento da Ideia (“Idea deployment”)**: acompanhamento da ideia como produto após a entrega ao público alvo. Que alimenta todas as etapas anteriores com o histórico do desenvolvimento respectivamente na forma de casos de sucesso para as fases “a” e “b”, estatísticas para a fase “c” e, por último, como padrões de projetos para a fase “d”.

Em cada abordagem, a participação dos usuários é reforçada pelo fluxo de informação e interação. Assim, o ciclo de vida organiza os aspectos relacionados a cada fase e o conteúdo gerado, portanto, em cada contexto, o número de etapas implica em um acúmulo de informação e estruturação das etapas próprio do processo. O processo cíclico é outra característica que difere a arquitetura de Westerski *et al.* (2011) com a de Perez *et al.* (2014). Devido a esse processo o sistema é retroalimentado e, portanto, não distingue início ou fim.

No modelo de ciclo de vida da ideia por Westerski *et al.* (2011), diferentes funções são exercidas em cada etapa (Figura 6). Como funções identificadas por Westerski *et al.* (2011) estão: clientes, comunidade, empregados, moderadores, revisores tomadores de decisão, especialistas do domínio, times de desenvolvimento, gestor de projeto, subcontratados e as empresas parceiras. Há ainda dois papéis muito específicos desse ciclo que são os organizadores da campanha de ideias e os administradores da gestão de ideias. O suporte de duas funções dedicadas ao ciclo demonstra a importância da gestão, do acompanhamento e da valorização das campanhas de geração de ideias.

Figura 6 – Ciclo de vida da ideia.



Fonte: Traduzida de Westerski *et al.* (2011).

2.2.2 Os Aspectos da Gestão do Conhecimento no ciclo

A participação intensiva de diferentes usuários e ainda os dados da organização geram um acúmulo de conhecimento que necessita ser gerenciado.

A relevância da gestão do conhecimento nos IMS é percebida em 15 artigos analisados – (Perez *et al.*, 2014); (Elerud-Tryde e Hooge, 2014); (Bettoni *et al.*, 2013); (Westerski *et al.*, 2013); (Poveda *et al.*, 2012); (Westerski *et al.*, 2012); (Spencer, 2012); (Westerski e Iglesias, 2011); (Westerski *et al.*, 2010); (Bettoni *et al.*, 2010); (Bailey e Horvitz, 2010); (Hrastinski *et al.*, 2010); (Richter e Vogel, 2009); (Paukkeri e Kotro, 2009); (Baloh *et al.*, 2007).

Perez *et al.* (2014) e Westerski e Iglesias (2011) identificam os IMSs como um sistema baseado em conhecimento, porém cada autor apresenta um foco nessa inserção.

Perez *et al.* (2014) defendem a necessidade de uma ligação com o conhecimento dos demais sistemas legados. A ligação com a documentação e sistemas mostra uma grande importância na agregação e alinhamento das estratégias organizacionais. Desse modo, um IMS desempenha um potencial ainda maior e diretamente ligado à tomada de decisões. Perez *et al.* (2014) percebe esta ligação ao menos de duas maneiras: a) em módulos extras de um sistema já utilizado; ou b) em um sistema a parte ligado aos demais sistemas.

Westerski e Iglesias (2011) definem IMS diretamente como um tipo de sistema de gestão do conhecimento, aplicado nas organizações para reunir a entrada de grandes comunidades. Como auxílio ao processo de inovação, Paukkeri e Kotro (2009) percebem nesses sistemas ainda a necessidade de integrar: a memória organizacional; e o conhecimento acumulado da empresa, por seus parceiros e em documentos.

Bettoni *et al.* (2010) também citam a "*gestão de ideias como um instrumento da gestão do conhecimento*". Alexander (2008) reforça esse conceito citando que "*o conhecimento é feito de ideias que são geradas, testadas e regeneradas por indivíduos e comunidades*", portanto essencial para o processo de gestão do conhecimento de toda a organização.

Essa gestão é identificada por Bettoni *et al.* (2010) nos processos de compartilhamento e aprendizagem em rede que ocorrem no sistema de gestão de ideias, pois representam atividades intensivas em conhecimento.

Portanto, além do conhecimento de sistemas ou documentos, é na interação e no enriquecimento de ideias que a participação da comunidade confere um potencial imprescindível ao processo. Estes conceitos reforçam a importância do paradigma de inovação aberta, aproximando clientes, parceiros, fornecedores e até os próprios colaboradores ao processo.

Campanhas bem-sucedidas nessas plataformas podem acumular rapidamente uma infinidade de comentários e publicações devido à “*crowd force*” (Murah *et al.*, 2013; Elerud-Tryde e Hooge, 2014). Logo, o desafio passa a ser o gerenciamento e tratamento adequado do conteúdo gerado. Porém a característica de dados não estruturado torna este processo complexo e custoso, o que despenderia um imenso tempo e esforço de ao menos “conhecedores” do domínio. Alternativas computacionais para esta tarefa são discutidas no tópico a seguir.

2.2.3 A representação do conhecimento em sistemas Web

A conexão de informações é a base da Web. Sendo um dos desafios agregar semântica a conjuntos de dados não estruturados do grande volume de dados e, assim, extrair o conhecimento. A capacidade por tornar esses dados compreensíveis e compartilháveis por diferentes agentes faz parte dos esforços da Web Semântica, sugerindo potenciais serviços de conexão de conhecimento (Gruber, 2008).

A Web Semântica é o nome que abrange o projeto de Tim Berners-Lee e da W3C (*World Wide Web Consortium*) ao qual pretende ser uma extensão da Web atual. A finalidade é permitir uma maior interação entre agentes humanos e computacionais, para ampliar a forma com que programas podem interagir com as páginas Web e também oferecer um uso mais intuitivo por parte dos usuários. Esta mudança transforma a Web da Informação na Web do conhecimento através da semântica aplicada com informações distribuídas e interconectadas (Berners-Lee *et al.*, 2001).

Nessa Web 3.0, os vocábulos controlados das Ontologias tomaram grande importância. A Ontologia, na área de ciências da computação, é definida como “*uma especificação formal e explícita de uma conceitualização compartilhada*” (Studer *et al.*, 1998). Considerada como formal por ser processável por máquina; explícita por definir todos os conceitos atributos e relacionamentos sem ambiguidade; conceitualização em um modelo conceitual de um domínio específico; e

compartilhada devido necessidade de um consenso entre especialistas do domínio.

A representação de conhecimento das ontologias segue o padrão de grafos de redes semânticas com a criação de sentenças RDF (*Resource Description Framework*) com sujeito, predicado e objeto (Berners-Lee *et al.*, 2001):

- Sujeito: um recurso ou coisa sobre o qual é feita uma declaração;
- Predicado: uma propriedade ou característica do sujeito que é expresso mediante a declaração;
- Objeto: valor da propriedade a qual se refere o predicado.

Nos Sistemas de Gestão de Ideias, as ontologias são utilizadas para a representação do conhecimento e semântica dos dados. Entre os padrões observados na revisão, a ontologia do Gi2MO¹ por Westerski *et al.* (2013) é a mais destacada.

Westerski *et al.* (2010) demonstram como modelar os dados de um IMS usando os princípios da Web Semântica descrevendo a pesquisa e as dificuldades que surgem no processo de concepção baseado em uma ontologia de domínio.

Segundo Westerski *et al.* (2013): “*A ontologia estabelece bases para a gestão do conhecimento com base na interligação de sistemas corporativos e ativos da Internet com o princípio de aumentar a consciência da informação e ajudar na avaliação da inovação*”.

Ligando dados de um Sistema de Gestão de Recursos Humanos (HRMS) ao IMS, Westerski e Iglesias (2011) relacionam ideias e o conhecimento de colaboradores. O autor descreve alguns objetivos dessa ligação como:

- Avaliar ideias com base nas competências, experiências e habilidades do criador;
- Recomendar colaboradores com base em suas habilidades e a relação dessas com a temática das ideias;
- Julgar a eficiência dos colaboradores ou revisores com base em sua atividade no IMS e projetos regulares da empresa;

¹ Disponível em <<http://www.gi2mo.org/ontology/>>, acesso em 16 de novembro de 2015.

- Promover usuários que são mais ativos em determinada área ou analisar a participação nos esforços de inovação de algum grupo.

2.3 AS ATIVIDADES DE AVALIAÇÃO E SELEÇÃO EM IMS

A inovação radical cada vez mais faz parte da discussão nas grandes organizações. Devido ao macro ambiente mais exigente e em constante transformação e, dessa forma, *“projetos de inovação incremental já não são suficientes para garantir um crescimento contínuo no setor privado e manter altos níveis de serviço no setor público”* (Brix, 2015). Além disso, *“descobertas de novos negócios e mudanças do cenário econômico não são apenas importantes para o crescimento, mas também são fundamentais para a sobrevivência”* (Brix, 2015) e, reconhecendo isso, as organizações enfatizam a inovação radical em sua intenção estratégica.

Porém, apesar de defenderem a importância de ideias novas e radicais de projetos, a maioria dos tomadores de decisão é relutante em alocar recursos humanos e financeiros necessários para a exploração desses tipos de propostas. Para Brix (2015), essa intimidação se deve aos níveis de incerteza e a orientação pelo benefício ao longo prazo de tais projetos radicais, o que pode não agregar valor imediato para a organização.

Esse paradoxo é o tema que Brix (2015) explora definindo como desvantagem evidente a utilização inadequada de ferramentas para orientar decisões, quando se trata de exploração de novas potencialidades proporcionadas pelo FEL. Sufocando a inovação por definições de "bom" ou "ruim" ao invés de moderar critérios de avaliação, e, assim aprender com o processo (Brix, 2015).

Os estudos de Brix (2015) defendem a necessidade de uma avaliação mais flexível também identificada por Sandström e Björk (2010). Para os autores, existe um risco evidente de que as ideias que são descontínuas sejam filtradas, não por serem consideradas ruins, mas porque *“não se encaixam no modelo de negócios atual e processo de avaliação”* (Sandström e Björk, 2010). Portanto, é necessário diferenciar a abordagem também de acordo com a natureza da inovação, pois *“empregar procedimentos de avaliação padrão podem ser ineficazes uma vez que inovações descontínuas são fundamentalmente diferentes das incrementais”* (Sandström e Björk, 2010).

O comentário vai de encontro com o conceito de “*falsos negativos*” descritos por Chesbrough (2003) ao qual conclui que as organizações necessitam alterar continuamente suas métricas de inovação, pois o alcance e as perspectivas sobre a inovação foram modificados a partir da preocupação com a inovação incremental de produtos, no sentido de olhar tanto para a inovação descontínua como também para o modelo de negócio da inovação e da inovação aberta.

Nos tópicos a seguir são apresentadas algumas técnicas apresentadas pelos autores em relação à avaliação e seleção de ideias.

2.3.1 Métodos e técnicas para a avaliação e seleção

A gestão de ideias em meio a um grande sistema frequentemente assume a forma de: “*Eu não sei exatamente o que eu quero, porém eu sei o que quando eu vejo isso.*” (Spencer, 2012)

Segundo Westerski e Iglesias (2011), separar boas e más ideias é a principal razão para a existência da gestão de ideias, e também “*uma das etapas mais importantes e problemáticas*”. Para executar esta tarefa de avaliação, colaboradores preenchem formulários que são os meios para comparação normalizada, filtragem e, finalmente, a seleção das melhores candidatas para a implementação.

As métricas geradas automaticamente, na maioria dos casos, se limitam a simples estatísticas derivadas da atividade da comunidade como: número médio de observações em tempo por ideia, por usuário; quantidade de comentários em um período; ranking de votos; entre outros. Baez e Convertino (2012) também identificam que as participações nos processos de avaliação utilizam as métricas da comunidade como número de comentários, votos ou importância. Esses números podem ganhar destaque nos estudos em redes sociais sobre o alcance orgânico, viral ou gratificado.

Westerski *et al.* (2011) apresentam três técnicas para lidar melhor com essa etapa que são:

- a) **A avaliação de ideia:** por comentários realizados periodicamente, em paralelo com o processo de seleção;
- b) **Pré-processamento de dados auxiliado por máquina:** tarefas pesadas computacionais como estatística, detecção de padrões, etc.;
- c) **Filtragem e armazenamento em clusters:** por métodos textuais e gráficos aplicados durante a seleção para aumentar a navegação e pesquisa;

Por fim, o enriquecimento com dados da revisão interna e da avaliação automática com ranking e triagem de ideias auxiliam na etapa de seleção. Segundo Westerski *et al.* (2011), a avaliação por colaboradores da organização serve para enriquecer a descrição criada pela comunidade. É nessa etapa que ocorre o alinhamento da ideia com as estratégias da organização, objetivos ou necessidades atuais. Para reunir essa entrada podem ser utilizadas ferramentas tais como: a) ferramentas de classificação; b) categorização; c) interligação de comentários; e d) revisões textuais.

Elerud-Tryde e Hooge (2014) organizaram um processo de triagem de ideias utilizando diferentes funções e grupos. Além disso, aplicaram diferentes critérios de avaliação para ideias categorizadas como radicais ou incrementais.

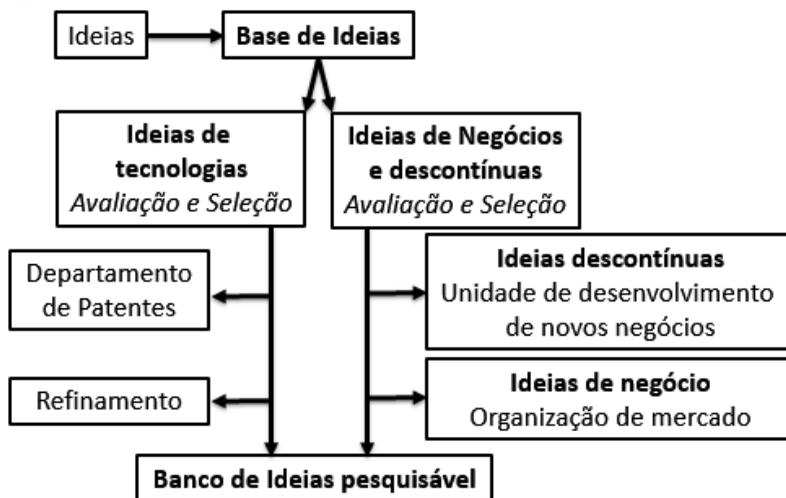
Alguns sistemas utilizam ainda a função de revisores, avaliadores ou facilitadores. Nessas funções usuários avançados, administradores ou gerentes ajudam a revisar, organizar e selecionar as ideias como elas são geradas, segundo a pesquisa de Baez e Convertino (2012) no contexto social e comunitário. Para os autores, o objetivo dessa divisão de papéis é facilitar o processo de seleção e julgamento em torno de boas ideias. A partir disso, então, ajudar as melhores ideias a surgir e crescer, comparando os comentários da comunidade com os dos gestores da organização ou líderes comunitários.

Em relação à avaliação por equipes, Bettoni *et al.* (2013) utilizam o Método *Solution Finder Model* (SFM) para avaliação das ideias. A partir da necessidade em comparar diferentes soluções como parte da avaliação e colocar em algum tipo de ordenação. Bettoni *et al.* (2013) identificam o interesse em uma "generalização do conhecimento" no sentido de um *framework* ao qual todos os critérios a serem avaliados sejam interpretados como "*especificações de um mesmo paradigma genérico*". Assim, apresenta o modelo estrutural SFM. O modelo considera que as contribuições possam ser "*interpretadas de maneira viável como afirmações diretas ou indiretas, quer sobre as necessidades ou acerca de objetivos ou soluções*".

Os autores Sandström e Björk (2010) também apresentam uma alternativa, ao qual, semanalmente, uma equipe de avaliação se reúne e compartilham suas discussões, após isso discutem o grau da ideia em uma escala de 1 a 5. Os critérios de avaliação são definidos a partir do valor da ideia para a empresa em termos de novidade e utilidade, sendo: 1 ponto - A ideia é bem conhecida e / ou difíceis de implementar; 2 pontos - A ideia é conhecida e / ou tem vantagens menores; 3 pontos - a

ideia é nova e útil; 4 pontos - A ideia é nova e tem uma atividade descoberta clara; ou 5 pontos - a ideia é nova e excelente. Nessa proposta, ideias que anteriormente eram consideradas indesejáveis podem ser mais desenvolvidas como um novo negócio (Figura 7).

Figura 7 – A seleção de ideias em um Sistema de Gestão de Ideias.



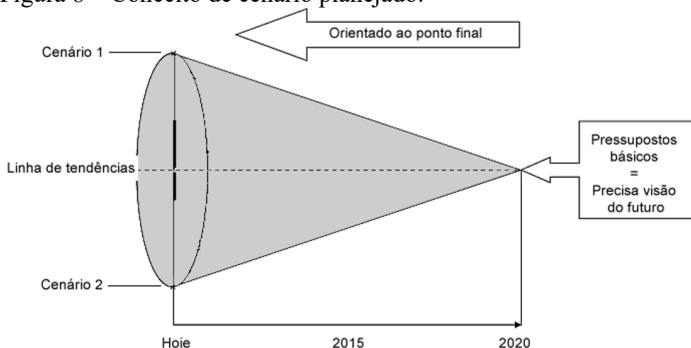
Fonte: Traduzido de Sandström e Björk (2010).

2.3.2 O tratamento de ideias rejeitadas

O tratamento para ideias está intimamente relacionado ao momento adequado e, no caso de uma avaliação, de quem julgará. Brem e Voigt (2009) mostram essa afirmação a partir de um funil que representa os cenários e uma linha horizontal que perpassa até o fim, sendo considerada as tendências. Ao fim do funil, uma projeção futura é representada pelas suposições básicas no planejamento (Figura 8).

Essa é a percepção retratada no momento de uma avaliação, o que definirá o aceite ou o descarte de uma ideia, mantendo o julgamento mesmo em mudanças do contexto.

Figura 8 – Conceito de cenário planejado.



Fonte: Traduzido de Brem e Voigt (2009), página 363.

Diversas características podem explicar o motivo da rejeição de uma ideia. Segundo Brix (2015), muitas propostas são rejeitadas de imediato, pois o promotor não consegue destacar os valores e/ou benefícios imediatos para a organização. Assim, como na discussão de projetos, em que muito do conhecimento criado e compartilhado é inadequadamente documentado e perdido, como cita Karlsson e Törlind (2013). A exceção são os casos que geram patentes, mesmo que não utilizadas em produtos finais, porém servem como ponto para ressuscitar ideias em um novo momento mais adequado.

Segundo Bettoni *et al.* (2010) todas as ideias devem ser armazenadas em um "banco de ideias". O banco de ideias é útil, pois "*as empresas devem garantir que ideias rejeitadas sejam reenviadas automaticamente assim que as condições externas mudem, por muitas vezes ditadas pelas influências do mercado e ambientes de negócios individuais*" (Gausemeier e Berger, 2004). Caso contrário, colaboradores podem sentir que a sua ideia é imediatamente rejeitada devido exclusões "*top-down*" (Bettoni *et al.*, 2010), deteriorando o sentido de colaboração e de sua identidade de participação no grupo.

Bettoni *et al.* (2010) também citam algumas características das ideias rejeitadas como: o conteúdo, ou seja, uma ideia é impraticável por razões técnicas; ou devido ao contexto organizacional, como por exemplo problemas de orçamento, prioridade de outras ideias, política organizacional ou concorrência.

Outras razões identificadas por Bettoni *et al.* (2010) para o rejeite de uma ideia estão relacionadas: (a) ao momento que é errado; (b) a ideia não está suficientemente desenvolvida e precisa de mais trabalho; (c) que os recursos não estão disponíveis; ou (d) que a ideia não

combina com as prioridades atuais; que eles se originam de fontes diferentes e também têm diferentes estágios de maturidade ou que a ideia não se encaixa no quadro existente de referência na empresa em questão.

Portanto, é necessário identificar as incertezas que compõem as avaliações. Afinal, o potencial estratégico e a viabilidade das propostas que são identificadas e conceituadas durante o processo da inovação radical representam um repositório de conhecimento riquíssimo e único para melhorar o desempenho global das características dos produtos.

Como forma de identificar tipos de informações que são acessíveis na base de dados a partir do IMS, Brix (2015) define três conceitos que identificam o nível de incertezas e são representados por cores, sendo:

- Conceito "verde" é muito específico sendo representado por um plano de ação detalhado e exato para implementação ou realização de uma "prova de conceito" ao tomador de decisões
- Conceito "azul" que é algo mais difícil de descrever, porém nem todos os elementos são incertos, assim a descrição torna-se menos específica ou mensurável, consequentemente um plano de negócio poderá ser estimado, mas não validado.
- Conceito "vermelho" é identificado por um elevado nível de incertezas, poucos elementos são exatos e os demais são baseados em uma previsão.

Para Brix (2015), nos casos de alta incerteza, há um potencial avanço dos conceitos se ainda houver elementos que podem ser melhor desenvolvidos. Como maneira de verificar o potencial de negócios o autor identifica ainda ações operacionais na redução das incertezas como: Planejamento de descoberta direcionado (Christensen e Raynor, 2013) ou Planejamento de Aprendizagem (Rice *et al.*, 2008).

2.4 REDES SEMÂNTICAS

A representação do conhecimento, para o raciocínio computacional, requer esquemas estudados pela Inteligência Artificial (IA). Dentre a ramificação da Inteligência Artificial em conexionista e simbólica, Heinzle (2011) cita como principais esquemas da IA

Simbólica: as Regras de Produção, os quadros e roteiros (ou *frames* e *scripts*), a lógica proposicional e dos predicados, e as redes semânticas.

A partir do foco da presente pesquisa e devido à utilidade na Web, as redes semânticas são a abordagem tomada para a representação do conhecimento. As redes semânticas se constituem em “*um conjunto heterogêneo de formalismos que têm em comum o uso de uma estrutura de rede formada por nós interconectados através de arcos rotulados, formando assim um grafo rotulado direcionado*” (Heinzle, 2011). A representação do conhecimento nessas redes aborda os nós como objetos, conceitos, situações ou ações; e as arestas como relações entre esses nós.

Em relação às redes semânticas na área de ciências da informação, Moreiro González (2015) descreve três modos de aplicações utilizados pela inteligência organizacional para a indexação e qualificação de recursos:

- a) **Indexação por popularidade:** quantidade de comentários, números de votos e destaques pela comunidade;
- b) **Qualificação da opinião dos leitores:** qualificação a partir dos comentários e análise de sentimento;
- c) **Indicação ou indexação mediante atribuição social de etiquetas:** análise das palavras chaves, folksonomias.

Na revisão do estado da arte foram observadas as abordagens de: indexação por popularidade, como no IMS apresentado por Westerski e Iglesias (2011); e, também, a qualificação da opinião dos leitores, como descrito por Baez e Convertino (2012). Porém, em nenhum dos IMS e das técnicas de avaliação foi abordada a análise das atribuições sociais como método de avaliação.

Logo, uma oportunidade a partir da abordagem apresentada por Moreiro González (2015) é percebida para contribuição da atividade de avaliação de ideias, desde o apoio para a definição de critérios.

Nas seções seguintes são discutidas as atribuições sociais de etiquetas a partir das características das redes semânticas na Web.

2.4.1 Semântica na Web

Apesar do foco da Web semântica na representação e extração do conhecimento, Mika (2007) retrata as lacunas da literatura em relação a visão e os detalhes de implementação de arquiteturas semânticas para a

Web, afinal, “*é difícil, se não impossível obter conhecimento que é formal, mas também dinâmico e amplamente compartilhado*” (Mika, 2007).

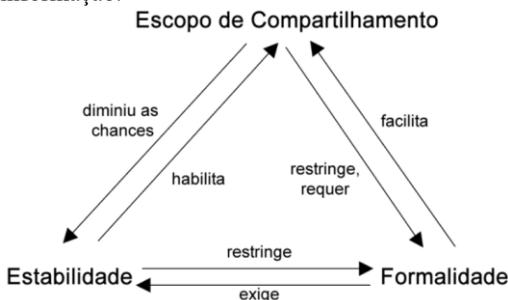
Contrário aos aspectos de um "contrato fixo e limitado" para a Web, Mika (2007) se baseia na frustração das ontologias em estabelecer um padrão capaz de representar a dinâmica dos dados na Web. Mesmo em relação à proposta das ontologias emergentes, ao qual diversos modelos de cooperação entre agentes autônomos racionais foram estudados com o objetivo de organizar o conhecimento, sugerindo então uma maneira mais escalável e de fácil manutenção para o desenvolvimento. Porém, essa visão de auto-organização continua sem soluções satisfatórias.

Van Elst e Abecker (2002) propõem uma classificação para gestão de informação em três dimensões: formalidade, estabilidade e escopo de compartilhamento do conhecimento, conforme observado na Figura 9.

Em determinado ponto o conhecimento – altamente formal – requer um compromisso que é difícil de alcançar em grande escala, enquanto ontologias – amplamente compartilhadas – são necessariamente rasas. A estabilidade é percebida, por exemplo, na Web devido a sua dinâmica, afinal, ao alcançar a formalidade, exigiria a estabilidade.

Dentro das dimensões de formalidade e compartilhamento de Van Elst e Abecker (2002), Mika (2007) destacou a Gestão do Conhecimento baseada em ontologias. As aplicações baseadas em ontologias são então classificadas de acordo com seu nível de expressividade (dimensão semântica) e de distribuição (dimensão da Web).

Figura 9 – Escopo de compartilhamento, Estabilidade e Formalidade da informação.



Fonte: Traduzido de Van Elst e Abecker (2002).

O próprio autor defende então que “*a Web Semântica é uma Web para máquinas, mas o processo de criação e manutenção é um processo social*” (Mika, 2007).

Ao estudar as relações em hipergrafos, Mika relata que: “*Há uma mágica clara como nós vemos semântica emergirem através das ações individuais de uma comunidade*”, percebe então a possibilidade de extrair semântica pertinentes a partir das subcomunidades da rede entre usuários. A formalização dessa visão se dá a partir dos estudos das folksonomias.

2.4.2 Folksonomia

O resultado da colaboração social gera um conjunto de marcações conhecidas como folksonomia. O termo “Folksonomia”, do inglês *folks* e *taxonomy*, é um neologismo para uma prática de categorização colaborativa usando palavras-chave (*tags*) livremente escolhidos pelos usuários (Mika, 2007).

Tsui *et al.* (2010) definem a folksonomia como o produto final de um processo pelo qual muitos usuários adicionam metadados na forma de palavras-chave ou etiquetas para compartilhar conteúdo. As folksonomias também são conhecidas como: etiquetagem colaborativa, classificação social, indexação social e marcação social.

A folksonomia permite aos usuários então “*descrever um conjunto de objetos compartilhados com um conjunto de palavras chave de sua própria escolha*” (Mika, 2007). Segundo Hotho (2010) Esse mecanismo de marcação social foi implementado em diversos ambientes de compartilhamento de conhecimento online desde que foi adotado pela primeira vez em 2004 pelo site de *bookmarking* social Delicious.

A associação de etiquetas de texto em recursos digitais por usuários é amplamente utilizada como método de indexação em sistemas. Essa ação de marcação pelos usuários é comum de ser encontrada em mídias sociais com o objetivo de criar metadados e ligações que poderão ser úteis para a classificação, navegação, recomendação ou recuperação de conteúdo (Lohmann e Díaz, 2012).

Em termos de representação do conhecimento, Mika (2007) cita que “*esse conjunto de palavras chave não pode sequer ser considerado como um vocabulário, a forma mais simples possível de uma ontologia na escala contínua de Smith e Welty (2001)*”, afinal o conjunto de palavras não é fixo.

Porém, esse conjunto de palavras forma um rico recurso para extração de conhecimento na Web. Ferreira *et al.* (2013) cita a utilidade dos vocabulários informais para “*desenvolvedores de sistemas e engenheiros de conhecimento na construção de modelos de domínio e em projetos que sofrem com o gargalo de aquisição de conhecimentos principalmente no que diz respeito ao número elevado de fontes de informação e divergências entre atores*”.

Portanto, o aspecto social e amplificado é o que diferencia a folksonomia da taxonomia. Enquanto a taxonomia classifica a partir de aspectos específicos, a folksonomia se baseia na avaliação (*rating*) e na análise a partir das palavras-chave marcadas (Lohmann e Díaz, 2012).

2.4.3 Conhecimento coletivo na Web

As tecnologias de inteligência artificial propiciaram a construção de sistemas inteligentes que agem de forma competente como indivíduos, incorporando a sua resolução de problemas. Porém é a aquisição do conhecimento que apresenta o principal limitante para esses sistemas inteligentes, em especial, devido à grande dificuldade em se obter o conhecimento de uma forma que as máquinas possam tirar proveito para resolver problemas reais.

Ou seja, as técnicas de aprendizagem de máquina e mineração de texto podem encontrar estruturas e padrões em grandes conjuntos, e colaborar no uso e coleta de dados. Porém, tais técnicas necessitam dos dados para relevar novos conhecimentos.

Gruber (2008) sugere que a internet é a melhor oportunidade de chegarmos a uma inteligência coletiva. Um ambiente que máquinas e humanos operem com sinergia em diferentes papéis. As pessoas como produtores e consumidores, fonte de conhecimento, de problemas e interesses do mundo real. E as máquinas como facilitadoras, armazenando e recuperando dados, buscando e combinando dados para inferir conclusões lógicas e matemáticas.

Assim como as pessoas aprendem a partir da comunicação, e frequentemente criam novos conhecimentos no contexto de uma conversa. A internet possibilita isso a partir da colaboração de máquinas e humanos para criar novos conhecimento e aprenderem uns com os outros de forma mais eficaz (GRUBER, 2008).

Porém, para Gruber (2008) a Web Social (ou Web 2.0) ainda não representa o conhecimento coletivo, pois a “*popularidade é uma medida*

de qualidade e não uma medida de autenticidade”, por isso o autor sugere uma Web Semântica Social.

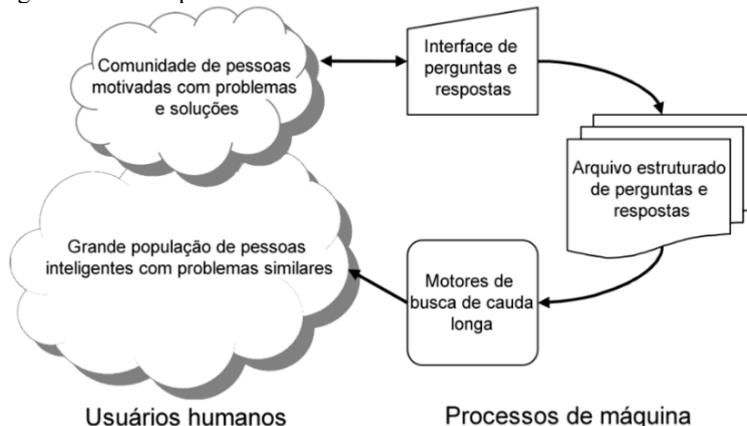
A Web Semântica Social (*Social Semantic Web*) une os padrões da Web Semântica e da Web Social a fim de permitir a criação de Sistemas de Conhecimento Coletivos (*Collective Knowledge Systems*) (GRUBER, 2008). Esta conexão permitiria então que, baseado nas interações dos usuários, seja compreensível em nível operacional a conexão dos conhecimentos, através de tarefas intensivas em conhecimento.

Somente a partir da composição dessas áreas que será possível tornar a observação dos indivíduos em um conhecimento coletivo, interpretável por máquina e compartilhável para uso em Sistemas de Conhecimento Coletivo (GRUBER, 2008).

A Figura 10 apresenta o exemplo de um Sistema de Conhecimento Coletivo proposto por Gruber (2008). O autor propõe que, nesses sistemas, agentes humanos e máquinas trabalham cooperativamente criando e organizando informações, interligando conteúdos e o conhecimento humano gerado.

Gruber (2008) denomina o exemplo de "FAQ-o-Sphere" que consiste em: (1) uma comunidade de colaboradores, participando em um processo social ampliado pela comunicação mediada por computador; (2) um motor de busca para a recuperação de informações, e (3) usuários inteligentes consultando ativamente o sistema, utilizando estratégias sintonizadas com o processo de geração de conteúdo e com o motor de busca, outros exemplos semelhantes são abordados com o mesmo viés.

Figura 10 – Exemplo de um Sistema de Conhecimento Coletivo.



Fonte: Traduzido de Gruber (2008).

Gruber (2008) ainda reforça sua expectativa mencionando que:

“[...] com o surgimento da Web Social, agora temos milhões de humanos oferecendo seus conhecimentos online, o que significa que a informação é armazenada, pesquisável, e facilmente compartilhada” e cita como desafio “encontrar a combinação certa entre o que é colocado online e métodos para fazer raciocínio útil para os dados”.

Tal pensamento vai de encontro com a proposta de Mika (2007). Mika (2007) discute as folksonomias como alternativa à aquisição de conhecimento. Enquanto as Ontologias representam o conhecimento por vocabulários controlados, as folksonomias representam o conhecimento a partir dos vocabulários populares. Segundo Mika (2007), a abordagem das ontologias formais se difere da abordagem probabilística, que ocorre na folksonomia, para o autor:

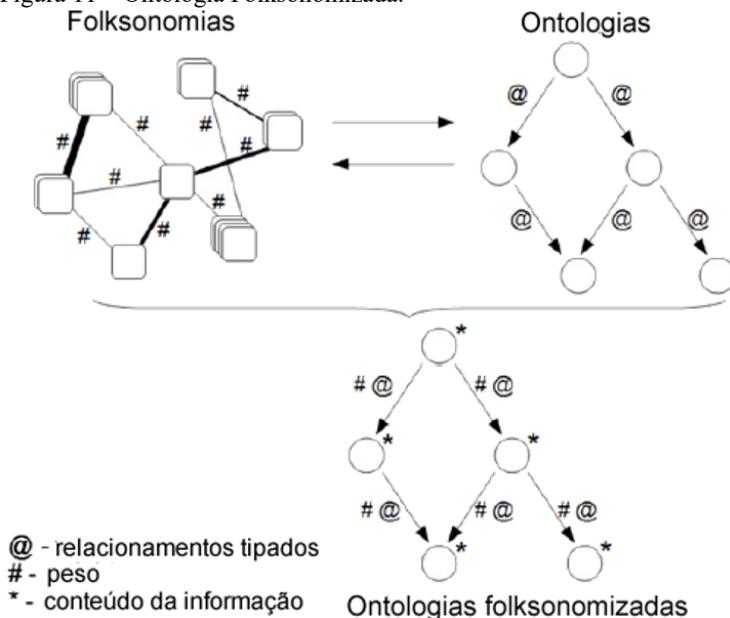
“[...] é a primeira vez que as barreiras de proporcionar conhecimentos foram reduzidas a tal ponto que os utilizadores comuns estão dispostos a fornecer os metadados sobre os recursos de Web em grande escala”.

Assim como Mika (2007), Gruber (2008) também defende que folksonomias e ontologias não são opostos, mas complementares. A partir dessa ideia, Gruber (2008) propõe a “*TagOntology*”, uma ontologia comum com foco na marcação social. Lohmann e Díaz (2012) também apresenta nove ontologias de marcação com aspectos distintos.

Já Alves e Santanchè (2013) defendem uma ontologia folksonomizada (*Folksonomized ontology*), ao qual o processo considera a extração, enriquecimento e evolução – “*3E steps technique*” – para a fusão das ontologias e da folksonomia (Figura 11).

Como foco para a pesquisa, será discutida a representação das folksonomias. Porém, tais pesquisas que relacionam a aquisição do conhecimento a partir das folksonomias juntamente com a formalização dos vocábulos por ontologias, apresentam uma importante tendência para trabalhos futuros.

Figura 11 – Ontologia Folksonomizada.



Fonte: Traduzido de Alves e Santanchè (2013).

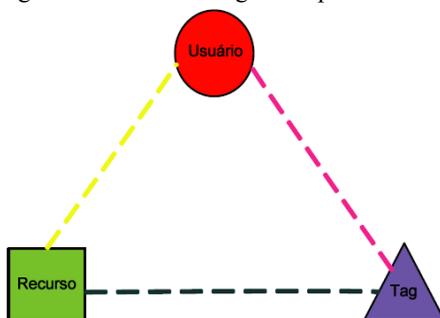
2.4.4 Representação das Folksonomias em Hipergrafos

A estrutura do vocabulário popular de uma folksonomia revela naturalmente um grafo a partir da relação dos usuários que marcaram recursos com etiquetas. Grafos são conjuntos de vértices e arestas ao qual cada vértice representa um objeto que pode conter atributos/propriedades e as arestas que conectam duas vértices. A notação de grafo é tal como: $G = (V, A)$.

Na marcação da folksonomia, o grafo é formado por três elementos centrais: usuários, recursos e *tags*. Os termos podem mudar para cada autor, porém o processo consiste em: um ou mais usuários (ou atores, pessoas, etc.) que marcam recursos (ou objetos, instâncias, etc.) com uma ou mais *tags* (ou palavras chave, marcações, etc.) (Lohmann e Díaz, 2012).

Zhang e Liu (2010) apresentam um modelo de grafo tripartido considerando a marcação social que relaciona as três entidades. O modelo em forma de triângulo permite analisar dinamicamente propriedades da rede a partir da marcação do usuário.

Figura 12 – Modelo de grafo tripartido.



Fonte: Traduzido de Zhang e Liu (2010).

Lohmann e Díaz (2012) apresentam outras representações formais de marcação e folksonomias em grafos sendo definidas em:

- a) **Modelo em três camadas:** Um modelo em três camadas de um grafo com arestas não direcionadas. O grafo representa relações binárias conectando usuários (U) a recursos (R) através de uma ou mais *tags* (T) (Figura 13a). Porém, esse modelo perde as informações sobre qual usuário anotou quais recursos. A representação é definida tal como: $A_{2 \times 2} \subseteq (R \times T) \cup (T \times U)$.
- b) **Modelo triplo uniforme:** Representa as relações entre usuários, recursos e marcações, porém a expressão de qual *tag* tem sido associada com qual usuário ainda é ambígua (Figura 13b). A representação da associação é definida tal como: $A_{3 \times 2} \subseteq (R \times T) \cup (T \times U) \cup (U \times R)$.
- c) **Modelo triplo não uniforme:** Relação ternária conectando todos os três elementos do conjunto, ao qual cada conjunto de três relações binárias é substituído por uma relação ternária, resultando em um hipergrafo triplo uniforme. Nesse modelo é possível considerar que determinado usuário marcou tal etiqueta para um recurso específico (Figura 13c). O conjunto de todas as anotações é então definido como $A_{1 \times 3} \subseteq R \times T \times U$.
- d) **Modelo de hipergrafo não uniforme:** Permite a relação n-ária conectando mais do que uma *tag*. Embora o modelo anterior já forneça uma representação simples da folksonomia, o hipergrafo é ainda conceitualmente impreciso, pois marcações que precisam de mais de uma etiqueta necessitam ser divididas em várias relações.

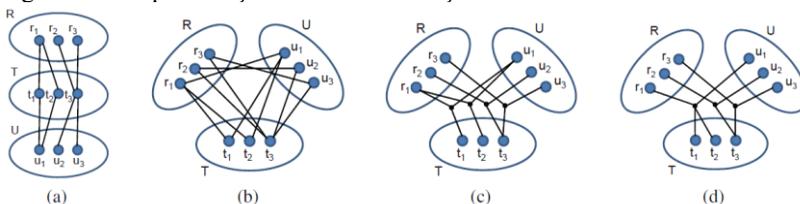
Logo, a representação conceitual mais precisa é um hipergrafo não uniforme com relações n -árias (Figura 13d).

Segundo Lohmann e Díaz (2012), com o objetivo de modelar as folksonomias como uma rede em um nível abstrato, Mika (2007) é o primeiro a representar o sistema como um grafo tripartido (três modos) com hiperarestas (*tripartite graph hyperedges*). Nesse grafo o conjunto de vértices é dividido em três conjuntos disjuntos tal como: $U = \{u_1, \dots, u_{kj}\}$, $T = \{t_1, \dots, t_{ij}\}$, $R = \{r_1, \dots, r_m\}$.

Esses conjuntos correspondem respectivamente ao conjunto de usuários, *tags* (conceitos, palavras chave) e o conjunto de recursos (objetos marcados). Portanto, estende o modelo tradicional bipartido (dois modos) de conceito e instâncias das ontologias através da incorporação de atores ao modelo. Assim, em um sistema de marcação social são criadas associações ternárias, ao qual o usuário marca objetos com conceitos. Mika (2007) define então a folksonomia por um conjunto de marcações tal como: $A \subseteq U \times T \times R$.

A rede entre formada pelos elementos é representada naturalmente como um hipergrafo com arestas ternárias, onde cada aresta representa o fato de que um dado usuário (U) associou certa *tag* (T) com um determinado recurso (R). A partir disso, uma representação dos vértices e dos nós pode ser definida em um hipergrafo de folksonomia A como um hipergrafo tripartido, tal como: $H(A) = \langle V, E \rangle$ onde $V = U \cup T \cup R$, $E = \{\{u, t, r\} \mid (u, t, r) \in A\}$.

Figura 13 – Representação formal de marcações e folksonomias.



Fonte: Lohmann e Díaz (2012)

2.5 ANÁLISE DE FOLKSONOMIAS

A partir das interações sociais da Web 2.0, diversos estudos apresentam formas de aquisição do conhecimento com técnicas de mineração de dados (*data mining*). Segundo Fayyad *et al.* (1996), a

mineração de dados é o "o processo não trivial de identificar padrões válidos, anteriormente desconhecidos e potencialmente úteis".

As ontologias são abordadas nessas análises como um importante recurso para a mineração. Porém, em um contraponto com as ontologias, que apresentam um afunilamento na aquisição de conhecimento (Mika, 2007), Hotho (2010) evidencia a mineração em folksonomias.

Ao apresentar a mineração em folksonomias, definida como "Folksonomy Mining", Hotho (2010) defende como uma forma significativa de aquisição do conhecimento na Web.

A relação de usuários, recursos e etiquetas de textos que formam as folksonomias, são representações de redes semânticas colaborativas, assim, muito da análise das folksonomias emergem dos estudos de análise de redes. Nos tópicos a seguir são apresentados alguns desses estudos voltados a aquisição de conhecimento a partir das folksonomias.

2.5.1 Decomposição

Em um hipergrafo tripartido é possível identificar todas as informações das atribuições de uma folksonomia. Porém, é a partir da decomposição em subgrafos que uma visualização simplificada dessas informações é possível (Landia *et al.*, 2013). O processo de decomposição de um grafo tripartido é representado por Lohmann e Díaz (2012). A Figura 14 demonstra essa decomposição, onde: "R" são os recursos; "U" são os usuários; e "T" são as *tags*.

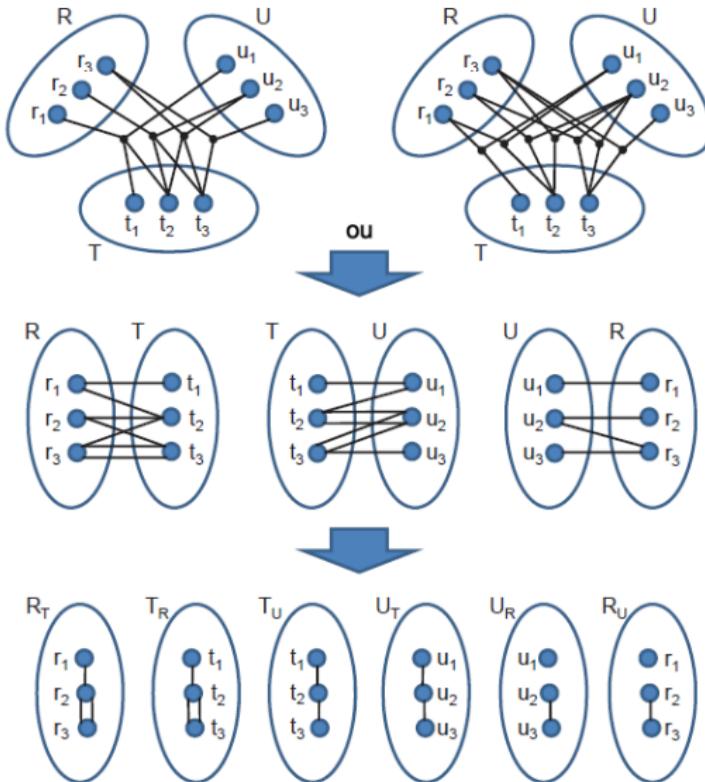
A partir dessa decomposição, Lohmann e Díaz (2012) apresenta como exemplo o conjunto das *tags*. Para esse caso, uma forma de transformação poderia considerar a quantidade que determinada *tag* coocorre para ser representada pelo tamanho do nó ou na distância entre eles. Outra forma seria considerar a frequência de cada *tag* para calcular um logaritmo e representá-lo pelo tamanho do nó ou da fonte.

Assim, essa possibilidade de redução é a transformação em grafos ponderados através da fusão das multiarestas e nessa transformação os pesos representam o número de ligações que existem entre os vértices, por vezes é necessário ainda determinar um limite para ocultar vértices, perspectiva similar a visualização de *tags* em nuvens (Lohmann e Díaz, 2012).

Assim, um grafo UT, por exemplo, pode ser definido tal como:

$$UT = \langle U \times T, E_{ut} \rangle, E_{ut} = \{(u, t) \mid \exists r \in R: (u, t, r) \in E\}, w: E \rightarrow N, \forall e = (u, t) \in E_{ut}, w(e) = |\{r: (u, t, r) \in E\}|.$$

Figura 14 – Decomposição completa de um hipergrafo.



Fonte: Traduzido de Lohmann e Díaz (2012).

Ao qual, cada ligação nesse grafo pode ser mensurada pela quantidade de vezes que cada usuário utilizou determinada *tag*, conhecido em análise de redes sociais como rede de afiliação (*affiliation network*), aonde esse peso determina a força de ligação. É possível ainda gerar dois grafos simples como uma rede de um modo, com as ligações dimensionadas pela similaridade entre os usuários e atribuições.

Para facilitar a compreensão do processo de dobragem de um grafo bipartido em uma rede de um modo, Mika (2007) define, então, uma matriz a partir do grafo tal como: $U = \{u_{ij}\}$, onde $u_{ij} = 1$ se o usuário está filiado com a *tag* t_j . Uma nova matriz é definida por: $S = \{s_{ij}\}$, onde $s_{ij} = \sum_{x=1}^k b_{ix}b_{xj}$.

No caso apresentado, as ligações, entre usuários que marcaram as mesmas *tags*, possuem pesos que mostram o número de *tags* que eles utilizaram em comum. Então, uma notação matricial de $S=BB'$, conhecida como matriz de filiação, permite definir uma rede com base nas filiações compartilhadas. Assim como uma matriz dupla $O=B'B$ pode apresentar um grafo semelhante mostrando então a associação de *tags*, ponderada pelo número de usuários que utilizaram ambas as *tags* nas marcações.

Nesses dois grafos, a diagonal das matrizes correspondentes contém a contagem de quantos conceitos ou pessoas uma determinada pessoa ou conceito foi afiliado com o grafo bipartido. Tais valores podem ainda ser normalizados, por exemplo, através do cálculo do coeficiente de Jaccard e novamente filtrados com base nos pesos relativos (Mika, 2007).

Portanto, um grafo UT, que apresenta a ligação de usuários e *tags*, pode ser dobrado em dois outros grafos. A análise desses grafos cria então uma rede social e uma rede semântica, porém o grafo original é que possui todas as informações para gerar as redes, não sendo possível reconstruí-lo (Mika, 2007). Outras possibilidades de análise dos grafos criados a partir de um hipergrafo tripartido são discutidos por Mika (2007), como os grafos TR e UR.

A partir de um grafo TR, outra rede semântica é observada, ao qual as ligações entre os conceitos são ponderadas pelo número de instâncias que estão marcadas por ambos os termos. Essa rede é semelhante ao método básico aplicado nos processos de mineração de texto, em que os termos são geralmente associados por suas coocorrências nos documentos (Lohmann e Díaz, 2012).

Já o grafo UR representa outra rede social, ao qual o peso da ligação de um par de usuários é dado pelo número de recursos marcados por ambos. Outra análise possível é a rede de recursos TR, com as associações que mostram o número de usuários que têm marcado determinada *tag* a um par de instâncias.

Portanto, tais associações apresentam oportunidades de redução do hipergrafo a fim de obter novas interpretações por intermédio dos subgrafos. Porém, tais subgrafos perdem informações que constam somente nos hipergrafos (Lohmann e Díaz, 2012). Assim, a decomposição é uma técnica fundamental para a mineração de folksonomias, mas não adequada para o armazenamento dos dados.

2.5.2 Classificação

As classificações são observações em relação ao destaque e tendências para os dados analisados pelo grafo. Ao classificar os nós a partir de propriedades ou estatísticas, é possível destacar elementos para a análise. Em relação a classificação de folksonomias, é possível encontrar variações de algoritmos comuns para a análise de dados (Landia *et al.*, 2013).

Xu *et al.* (2008) propõem a classificação das *tags* em três classes: (a) *Tags* pessoais, usadas por somente um usuário e marcadas em muitas páginas, úteis para buscas individuais, mas sem utilidade para a comunidade; (b) *Tags* não populares, marcadas para diferentes recursos, por diferentes usuários, mas uma ou poucas vezes; e (c) *Tags* populares, são marcadas para diversos recursos, frequentemente por diferentes usuários e elas podem ser tomadas como etiquetas globais, que são usadas por muitos usuários.

A partir dos estudos de Xu *et al.* (2008) e de outros autores, Cantador *et al.* (2012) propõem a categorização baseada nas etiquetas mais populares encontradas no serviço de compartilhamento de fotos Flickr®, o autor categorizou em: (a) Baseadas em conteúdo, sendo subcategorizada como entidade física ou não física; (b) Baseadas em contexto, como localização ou tempo; (c) Subjetivas, tais como opiniões ou qualidade; e (d) Organizacional, subcategorizada em auto referência, tarefas ou ações.

Hotho (2010) discute o algoritmo “*FolkRank*”, um algoritmo de classificação baseado em grafo e modelado com base no popular *PageRank* da Google (Page *et al.*, 1999). Nessa variação do algoritmo, basicamente, quando um usuário influente marca uma *tag* importante em um recurso, essa marcação simetricamente torna relevante o recurso, a *tag* e o usuário. Portanto, o algoritmo não representa uma medida de popularidade, mas uma medida de reputação.

Segundo Landia *et al.* (2013), a técnica é utilizada especialmente como método para detecção de tendências e alterações da folksonomia. Um exemplo prático seria verificar quais *tag* têm ganhado ou perdido popularidade em um determinado período de tempo. Tais análises também são abordadas por Ghoshal *et al.* (2009) para inferir predições dentro do domínio observado.

Em uma análise ainda mais minuciosa, Gedikli (2013) apresenta o “*Local ranking*”. Ao qual a análise classifica a importância local das instâncias, baseada em uma área menor do grafo, o que representa uma

consideração mais apurada para conceitos específicos e suas atribuições relacionadas.

Em relação a importância local, uma possibilidade discutida é o reconhecimento de autoridades nas folksonomias, um conceito definido como “*folkauthority*” por Côgo e Pereira (2014). O propósito desse conceito é reconhecer usuários da comunidade que são destaques em determinados assuntos ou categorizar fontes de informação.

Outras análises apresentadas por Landia *et al.* (2013) são: “*ContentFolkRank*”, inclui o conteúdo de documentos diretamente no grafo (usuário, documento, *tag*); e o “*SimFolkRank*”, ao invés de etiquetas, inclui informações dos documentos implicitamente nos nós, constam nessas informações um vetor com o cosseno de similaridade normalizado por TF-IDF (“*term frequency-inverse document frequency*”, que significa frequência do termo–inverso da frequência nos documentos).

2.5.3 Distribuição

Os estudos de agrupamentos e detecção de comunidades são discutidos como forma de analisar a distribuição dos nós no grafo. As técnicas se baseiam em perceber nos dados um agrupamento que gere subcomunidades com dimensões reduzidas e mais próximas para a análise, por tanto, um mundo menor no grafo completo.

Segundo Ghosh *et al.* (2011), em teoria dos grafos, os métodos de distribuição são similares ao problema de particionamento de dados, o objetivo é agrupar de tal forma que seja possível analisar subcomunidades no conjunto.

Os estudos de Cattuto *et al.* (2007) sugerem que as distribuições de probabilidades para *tag* e recursos exibem uma lei de potência pura, enquanto a distribuição para usuários caracteriza um comportamento diferente em pequenas frequências, portanto, um “pequeno mundo”. O autor analisa então a existência de classes de usuários, alguns muito mais ativos do que a maioria.

Cattuto *et al.* (2007) investigam esse comportamento a partir das características de comprimento do caminho e coeficiente de agrupamento em hipergrafos. A característica de comprimento do caminho (*path length*) em um grafo descreve o tamanho médio de um caminho mais curto entre dois nós aleatórios. Quanto menor o comprimento do caminho, em média, menos pulos são necessários para alcançar um nó a partir de outro nó.

Como folksonomias são estruturas triádicas de atribuições (*tag*, usuário e recurso), então um caminho é definido como: “*uma sequência de hiperarestas de tal forma que em cada hiperaresta ao menos o usuário ou o recurso ou a tag são compartilhados com a hiperaresta seguinte*” (Cattuto *et al.*, 2007). Os autores definem dois coeficientes de agrupamento como recomendação:

- a) **Exclusividade (*Cliquishness*)** – O coeficiente de agrupamento de um nó é alto se muitas das possíveis arestas estão dentro da vizinhança. Um alto coeficiente indicaria, por exemplo, que muitos dos usuários relacionados a um determinado recurso atribuíram conjuntos de *tag* sobre esse. O mesmo pode ser definido simetricamente para *tag* e usuários
- b) **Conectividade ou Transitividade** – Permite a noção de que o agrupamento em torno de um nó é alto se muitos nós da vizinhança do nó foram conectados entre si, mesmo se esse nó não está presente.

A fração de pares vizinhos que permaneceriam conectados se determinado nó fosse excluído, indica que os arredores dos recursos contêm combinações únicas de *tag* e usuários que só ocorrem uma vez. A partir dos conjuntos de dados, Cattuto *et al.* (2007) observam estruturas de usuários ligados constantemente a determinadas *tag* e definem isso como “personomias”. Uma espécie de folksonomia do usuário.

Já Jäschke *et al.* (2008) sugerem o algoritmo TRIAS que permite a descoberta de subconjuntos de usuários que implicitamente concordam sobre uma conceptualização. Esses usuários convergem um subconjunto de marcações em conceptualizações compartilhadas de recursos.

Porém, a heterogeneidade dos termos é apontada por Mika (2007) como uma desvantagem significativa dos vocabulários, pois “*termos tomam sentidos diversificados dependendo do contexto*”. Uma alternativa seria o que o autor denominou de “enriquecimento com semânticas adicionais”.

Esse enriquecimento consiste na análise da rede em que palavras gerais são mais propensas a se relacionar com diferentes grupos de palavras, enquanto termos mais específicos são esperados apresentarem um agrupamento mais denso com seus vizinhos (Mika, 2007). Portanto, o cálculo de coeficiente de agrupamento, a centralidade local (*betweenness centrality*) ou a restrição sobre os termos permitem distinguir os termos em geral ou específico (Mika, 2007).

Outra possibilidade percebida por Mika (2007) é que, a partir da mesma análise, é possível que algoritmos de agrupamento auxiliem a encontrar conjuntos de sinônimos dos termos mais específicos.

2.5.4 Recomendação

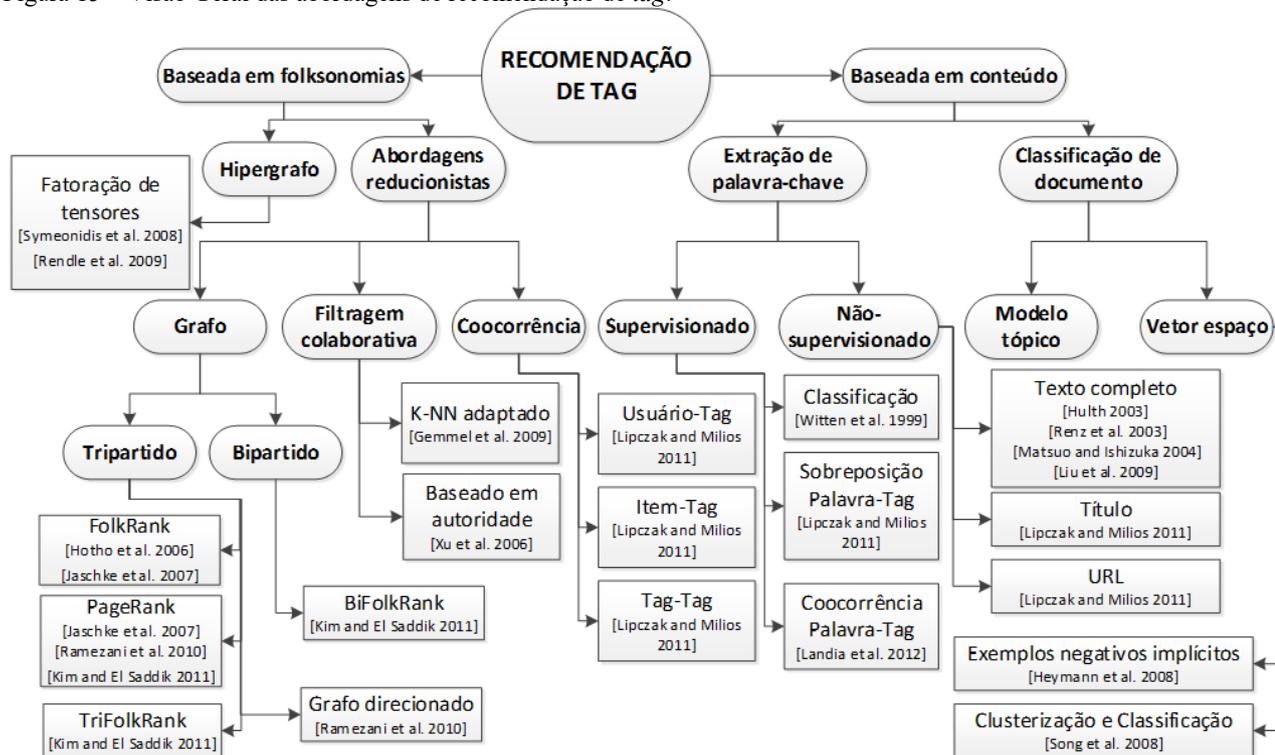
Burke (2002) define o termo “Sistema de Recomendação” como qualquer sistema que produza recomendações individualizadas ou que, de forma personalizada, conduzem o usuário para objetos úteis e interessantes em um espaço de opções possíveis.

Esses sistemas têm grande utilidade em um ambiente onde a quantidade de informações online supera amplamente a capacidade de qualquer indivíduo para examiná-la. Em um sistema de recomendação são os critérios de informação "individualizada" e "interessante e útil", que os distingue de sistemas de recuperação de informação ou motores de busca (Ricci *et al.*, 2011), dado que um sistema de recuperação apenas conduz o usuário a alguma informação disponível.

A recomendação também é amplamente utilizada no comércio eletrônico baseada no perfil e interação dos usuários, ou nas características dos itens, ou de forma híbrida. Já na recomendação baseada em conhecimento, o diferencial está na forma como as soluções são calculadas (Ricci *et al.*, 2011). Nessa abordagem baseada em conhecimento, a função de similaridade estima o quanto as necessidades ou a descrição do problema, coincidem com as recomendações ou soluções do problema.

Em relação aos Sistemas de recomendação para marcações sociais, Landia *et al.* (2013) apresenta uma visão geral por diferentes abordagens, conforme Figura 15.

Em cada abordagem dos sistemas de recomendação há pesquisas, algoritmos e cálculos para a análise. Nesse sentido, Gemmill *et al.* (2010) propõem que cada componente de uma recomendação em uma folksonomia pode ser definido com uma função tal como: $\psi : U \times R \times T \rightarrow R$, onde produz um resultado de valor real “p” como a relevância prevista de um elemento para um par de recursos: $\psi(u, r, t) = p$. O autor continua a discussão apresentando modelos de recomendação a partir de cálculos que resultam em “p”.

Figura 15 – Visão Geral das abordagens de recomendação de *tag*.

Fonte: Traduzido de Landia *et al.* (2013).

Tais cálculos de similaridade podem se valer de diferentes algoritmos como: SimRank, Teorema de Pitágoras, similaridade pessimista, cosseno de similaridade, ajuste do cosseno, coeficiente de Pearson, termo–inverso da frequência nos documentos (TF/IDF), discutidos por Xu *et al.* (2008).

Dentre esses, o coeficiente de correlação de Pearson é indicado para dados mensuráveis ou que tenham peso, pois resulta em uma medida linear para o cálculo da relação entre pares. Porém, segundo Xu *et al.* (2008), algumas dessas marcações precisam ser podadas para evitar resultados imprecisos. A poda, nesse caso, poderia ser feita envolvendo o uso de técnicas como: (a) número mínimo de votos em comum; ou (b) "*default voting*".

Em especial a técnica "*default voting*" é apresentada como uma extensão da correlação de Pearson por filtragem colaborativa, ao qual a similaridade entre os pares é calculada a partir do peso das classificações, em apenas interseções de itens que ambos usuários ou recursos foram marcados (Su e Khoshgoftaar, 2009).

2.5.5 As ferramentas de Análise

Uma das principais características dos IMS é a massiva contribuição da comunidade e a variação de informação por tempo. Porém as limitações das ontologias e processamento necessário se tornam pontos fracos no gerenciamento da informação.

Com uma aplicação semelhante Cattuto *et al.* (2013) abordam a alternativa de bancos de dados em grafo, o Neo4j. Banco de dados em grafos permitem uma gerência maior das informações, mantendo as redes semânticas, axiomas e armazenamento. Já o software Protégé, amplamente utilizado para estudos relacionados a tecnologias semânticas, agrega ferramentas de inferência, porém mais voltado para o desenvolvimento de ontologias (Tudorache *et al.*, 2013).

Mika (2007) cita como ferramentas para a análise de redes os softwares: UCINET, mais utilizado em ciências sociais para analisar dados de pesquisas sociométricas (Borgatti *et al.*, 2014); e o PAJEK, próprio para redes das mais variadas áreas com milhares de vértices e arestas (De Nooy *et al.*, 2011).

Uma opção mais simples é o NodeXL (*Network Overview Discovery Exploration for Excel*), um software baseado na planilha eletrônica Microsoft Office Excel® com o objetivo de ser "*uma ferramenta que evita o uso de linguagem de programação para as*

formas mais simples de manipulação de dados e visualização, para abrir a análise de rede para uma população mais ampla de usuários." (Quinn et al., 2012).

O Gephi é outra alternativa para a exploração e manipulação visual de redes disponibilizado em código aberto. Segundo Heymann e Le Grand (2013), o software se destaca dos demais no gerenciamento de redes que tenham propriedades nos nós, como as redes associadas às análises de *business intelligence* (BI). As propriedades são pares de valores-chave associadas a cada nó ou cada relacionamento. Por exemplo: uma ideia pode ter atributos como título e data de publicação. As ferramentas, módulos e estatísticas disponibilizadas permitem uma melhor e mais abrangente observação das características das redes (Bastian et al., 2009).

Tais ferramentas encontradas na literatura, entre outras opções, apresentam peculiaridade e diferenciais próprios da aplicação. Logo, a utilização e características das redes que indicaram a melhor opção ou agregação.

2.6 CONSIDERAÇÕES DO CAPÍTULO

Nessa pesquisa, a Inovação é percebida como um processo com o objetivo de gerar um conhecimento novo ou melhorado (Baregheh et al., 2009). Devido ao constante fluxo de informação, o processo necessita ser permeável aos atores externos, paradigma conhecido como Inovação Aberta (Chesbrough, 2003).

Koen et al. (2002) representa tal processo em três fases: (a) o *front end* da inovação; (b) o processo de desenvolvimento de novos produtos; e (c) a comercialização ou implementação. Sendo o *front end* considerado o momento mais complexo e custoso.

Durante o *front end*, uma potencial inovação é representada por uma simples e breve ideia, geralmente registrada rapidamente em uma base de ideias. Na internet, os sistemas adequados para tratar as bases de ideias são denominados Sistemas de Gestão de Ideias (IMS). Nesses sistemas, as ideias seguem um ciclo da geração, enriquecimento até que são avaliadas, selecionadas e consideradas aptas para a pesquisa e desenvolvimento.

O estado da arte apresenta modelos de avaliação que consideram a avaliação por especialistas ou pela comunidade. Para os especialistas, a organização define critérios a serem discutidos em formulários. No caso da avaliação pela comunidade são tomadas métricas em relação a

popularidade e, por vezes, comentários são explorados por processamento de máquina.

Se, de certa forma é complexo gerar critérios para avaliar ideias logo nas etapas iniciais, avaliações por métricas de engajamento social não representam o conhecimento aplicado, como cita Gruber (2008)

Logo, compreender o conhecimento aplicado pela comunidade para avaliar uma ideia, habilita que as organizações tenham métricas mais confiáveis para considerar e relacionar com seus reais objetivos. A representação desse conhecimento pode se valer das marcações sociais conhecidas como Folksonomias. As folksonomias representam o resultado da interação social ao qual usuários atribuem etiquetas livremente a recursos.

Os métodos para extrair semântica das folksonomias fazem parte da análise de rede. Uma das características apresentadas nessa análise é a mudança ao longo do tempo, ou seja, a dinâmica que ocorre em cada marcação, diferente das estruturas estáveis de vocabulários formais. Nessa dinâmica, partes tendem a se estabilizar ao longo do tempo, Mika (2007) compara esse período com as discussões e debates que ocorrem na construção das ontologias. Para o autor, essas partes podem ser então consideradas como ontologias leves que representam um consenso sobre o significado dos termos.

Hotho (2010) indica diversas análises para essas estruturas criadas pela folksonomias, entre os autores pesquisados, foram: (a) Decomposição de grafos; (b) Classificação; (c) Distribuição; e (d) Recomendação. O potencial dessas estatísticas e os recursos importantes para o tema de avaliação de ideias são definidos nos tópicos a seguir.

No próximo capítulo, essas referências são retomadas como base da fundamentação teórica do modelo apresentado. A bibliometria dos estudos é disponibilizada nos tópicos: “APÊNDICE B – Bibliometria de IMS” e “APÊNDICE C – Bibliometria de Folksonomias”.

3 PROCEDIMENTOS METODOLÓGICOS

Neste capítulo é apresentada a metodologia utilizada para o desenvolvimento da pesquisa. Na primeira seção é caracterizada a metodologia, seguida pelos procedimentos metodológicos.

Essa pesquisa é caracterizada pelo cunho tecnológico para solução de um problema, e, portanto, uma contribuição pragmática e, conseqüentemente, de impacto científico.

Na busca por uma metodologia adequada ao propósito na linha tecnológica da pesquisa de Engenharia do Conhecimento aplicada às organizações, Braga (2012) identifica 5 propostas de metodologia para a área de pesquisa aplicada: March e Smith (1995); CommonKADS por Schreiber (2000); Metodologia CISM (acrônimo de “*Composition, Environment, Structure, Mechanism*”) por Bunge (2003); Von Alan *et al.* (2004); e a *Design Science Research Methodology* (DSRM) por Peffers *et al.* (2007).

Nessa seção é apresentada a metodologia *Design Science Research* como modelo científico para o processo, após a apresentação da metodologia é discutida a relação dessas com a estrutura da dissertação, assim como os procedimentos utilizados, como a revisão integrativa discutida na seção posterior. O embasamento científico desta pesquisa é fundamentado nas orientações metodológicas descritas nas seções seguintes.

3.1 METODOLOGIA DE PESQUISA DA DESIGN SCIENCE

A ciência fatual, que engloba as ciências naturais e sociais, tem como objetivo explorar, descrever, explicar e prever fenômenos para desenvolver o conhecimento em determinadas áreas (Simon, 1996). Porém, por vezes, é necessário projetar, criar e construir artefatos aplicáveis para solucionar problemas, e portanto, não é suficiente apenas a descrição ou explicação de uma situação para o avanço do conhecimento (Van Aken e Romme, 2012).

Simon (1996) diferencia o artificial do que é natural, em “As Ciências do artificial”, originalmente de 1969. Para o autor, o artificial é algo produzido, inventado ou que sofre intervenção humana, logo a ciência do artificial tem a preocupação de solucionar problemas conhecidos ou projetar algo que ainda não existe. Assim Simon introduz o conceito de *Science of design*, denominada posteriormente de *design Science*. Relacionada fortemente com as áreas de estudo das ciências

aplicadas como engenharia, medicina, direito, arquitetura e educação, em especial a de sistemas de informação.

A *Design Science*, em português ciência do artificial ou ciência do projeto, busca projetar e produzir sistemas que ainda não existem e modificar situações existentes (Dresch *et al.*, 2015). Peffers *et al.* (2007) define os objetivos da *Design Science* em projetar, criar e avaliar artefatos de tecnologia da informação destinados a resolver problemas organizacionais identificados em um processo rigoroso a fim de resolver os problemas observados e então comunicar os resultados ao público interessado. Os resultados podem ser considerados em inovações sociais, novas propriedades técnicas, sociais e/ou recursos de informação (Peffers *et al.*, 2007).

Portanto, a pesquisa em *Design Science* formula, prescreve e valida soluções para alcançar melhores resultados com foco na solução de regras de *design* (Dresch *et al.*, 2015). Essa prescrição é a característica fundamental que posiciona a *Design Science* como um paradigma epistemológico, diferenciado pela prescrição facilitada para que os resultados sejam ensinados, aprendidos e colocados em prática pelos profissionais nas organizações, o que favorece a relevância prática da pesquisa (Peffers *et al.*, 2007).

O *Design* é então percebido por Dresch *et al.* (2015) como a realização de “*mudanças em um determinado sistema a fim de transformar as situações em busca da sua melhoria*”, efetuada pelo homem que, para tal fim, aplica o conhecimento para desenvolver artefatos que ainda não existem. Em resumo, o *design* compreende a concepção, projeto e concretização em circunstâncias definidas a serem aplicadas por profissionais quando pertinente. Já a *design science* é a base epistemológica quando se trata do estudo do que é artificial. E, por sua vez, a *design science research* é o método que fundamenta e operacionaliza a condução da pesquisa quando o objetivo a ser alcançado é um artefato ou uma prescrição.

A *design science research* como método de pesquisa orientado à solução de problemas busca: “*construir e avaliar artefatos que permitam transformar situações a partir do entendimento do problema, alterando suas condições para estados melhores ou desejáveis*” (Dresch *et al.*, 2015), portanto para diminuir o distanciamento entre teoria e prática.

Sua concepção consiste então na “*produção de um conhecimento científico envolvendo o desenvolvimento de uma contribuição científica, um conceito de solução que representa uma solução para uma gama maior de problemas*” (Peffers *et al.*, 2007). Esse conceito é então

avaliado em função de critérios relacionados à geração de valor ou utilidade. Em síntese, a DSRM “*inclui qualquer objeto projetado com uma solução incorporada a um problema de investigação compreendida*” (Peffers *et al.*, 2007).

3.1.1 As atividades do modelo

Baseado em outros sete modelos e no consenso de autores anteriores, Peffers *et al.* (2007) descreve uma síntese de seis atividades comuns apresentadas pela DSRM:

- **Atividade 1 – Identificação e motivação do problema:** Definir o problema específico da investigação e a justificativa da importância de sua solução alinhada com a relevância do problema. Justificar o valor de uma solução motiva o pesquisador e o público a alcançar a solução e admitir os seus resultados. Nessa atividade, conhecer o estado da arte do problema e a importância de sua solução é fundamental.
- **Atividade 2 – Definição dos objetivos de uma solução:** Inferir os objetivos para uma solução considerando o problema e o conhecimento do que é possível e viável de maneira quantitativa ou qualitativa. Para essa atividade é necessário conhecer as soluções atuais, caso existam, e suas características.
- **Atividade 3 – Projeto e desenvolvimento:** Projetar e criar o artefato, que pode ser identificado como construto, modelo, método ou instância de acordo com a abrangência definindo as funcionalidades desejadas, arquitetura e desenvolvimento. É considerada nessa atividade o conhecimento da teoria aplicada em uma solução.
- **Atividade 4 – Demonstração:** Demonstrar a aplicação do artefato para as instâncias do problema em experimentações, simulações, estudo de caso, ou atividade apropriada. Exige o conhecimento efetivo de como o artefato é utilizado para resolver o problema.
- **Atividade 5 – Avaliação:** Observar e estimar o quão bem o artefato suporta uma solução para o problema, comparando os objetivos com os resultados reais obtidos. Requer conhecimento de métricas relevantes e técnicas

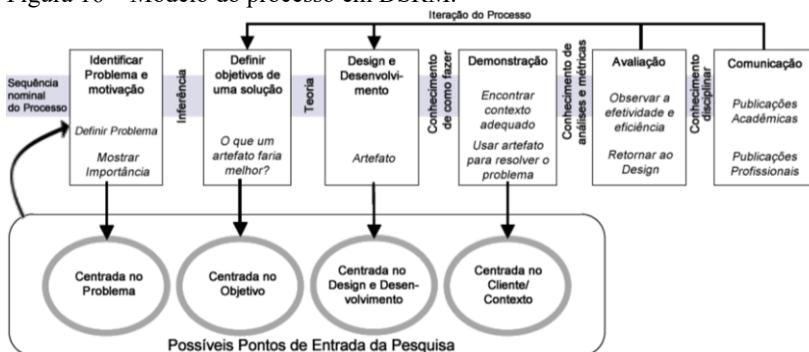
de análise. Se necessário, é possível voltar ao passo três para melhorar a eficácia do artefato, ou continuar e deixar detalhes para projetos futuros.

- **Atividade 6 – Comunicação:** Divulgar o problema e sua relevância, o artefato resultante, sua utilidade e ineditismo, o rigor da sua concepção, e sua eficácia para os investigadores e outros públicos relevantes, tais como a área profissional, quando apropriado. Essa etapa difunde o conhecimento resultante e para isso é necessário o conhecimento da cultura disciplinar.

O processo dispõe um modelo para que métodos sejam aplicados direcionando o projeto de pesquisa em um delineamento característico do rigor científico. Em resumo, entende-se por método um conjunto de passos utilizados para executar uma tarefa (March e Smith, 1995).

Os passos do método ocorrem sequencialmente com ações de *feedback* entre as etapas. A Figura 16 apresenta o modelo por Peffers *et al.* (2007).

Figura 16 – Modelo do processo em DSRM.



Fonte: Traduzido de Peffers *et al.* (2007).

3.1.2 Os pontos de entrada do modelo

Na DSRM, Peffers *et al.* (2007) identificam quatro diferentes pontos de entrada respectivamente sobre as quatro primeiras atividades do modelo, esses pontos são selecionados em função do contexto que a proposta da pesquisa, se origina:

- a) **Centrada no Problema:** Caso a ideia da pesquisa seja resultado da observação do problema ou a partir de pesquisas futuras sugeridas por um artigo científico ou projeto anterior.
- b) **Centrada no Objetivo:** Desencadeada por uma necessidade, industrial ou de uma pesquisa, que pode ser abordada através do desenvolvimento de um artefato
- c) **Centrada no Design e Desenvolvimento:** A partir da existência de um artefato que não tenha sido ainda formalmente pensado como uma solução para o domínio do problema explícito em que ele será aplicado. Tal artefato poderia vir de outro domínio de pesquisa, já utilizado para resolver um problema diferente, ou ter surgido como uma ideia analógica.
- d) **Centrada no Cliente ou Contexto:** Iniciada na observação de uma solução prática que funcionou, resultando em uma solução de *Design Science* se os pesquisadores regressarem o trabalho para aplicar o rigor do processo científico de maneira retroativa. Isso poderia ser o subproduto de uma experiência em consultoria.

3.1.3 Os artefatos resultantes

A *Design Science Research* é a abordagem metodológica que consiste na construção de artefatos que trazem benefícios às pessoas (Peffer et al., 2007). Assim o interesse do projeto está na definição do “o que” e “como” as coisas devem ser, portanto a concepção de artefatos que realizem determinados objetivos (Simon, 1996).

Dresch et al. (2015) definem um artefato como “a organização dos componentes do ambiente interno para atingir objetivos em um determinado ambiente externo”. A partir dessa definição os autores relacionam três elementos sendo determinados como: o propósito ou objetivo; o caráter do artefato; e o ambiente em que ele funciona.

Nessa perspectiva, a pesquisa em *design science* pode gerar cinco tipos de resultados científicos de acordo com a abrangência. Desses resultados, quatro são equivalentes com os propostos por March e Smith (1995) para pesquisas aplicadas ou naturais e o quinto uma proposição de Van Aken e Romme (2012) abordado por Dresch et al. (2015). A seguir são listados os artefatos:

- a) **Construtos:** Construtos ou conceitos formam o vocabulário de um domínio. Eles constituem uma conceituação utilizada para descrever os problemas dentro de um domínio e para especificar as respectivas soluções. Eles formam a linguagem especializada e conhecimento compartilhado de uma disciplina ou subdisciplina.
- b) **Modelo:** Um modelo é um conjunto de proposições ou declarações que expressam relações entre os construtos. Em atividades de design, os modelos representam situações como declarações de problema e solução.
- c) **Método:** O método é um conjunto de passos, ou uma orientação ou algoritmo, usado para executar uma tarefa. Métodos são baseados em um conjunto de construtos básicos (língua) e uma representação (modelo) do espaço de solução.
- d) **Instanciação:** Uma instanciação é a realização de um artefato em seu ambiente. Instanciações servem para demonstrar a viabilidade e a eficácia dos modelos e métodos que eles contêm. Instanciações operacionalizam construtos, modelos e métodos.
- e) **Design Propositions:** Esse tipo de artefato é ainda mais abrangente e abordado por (Van Aken e Romme, 2012) como um padrão genérico que pode ser utilizado para o desenvolvimento de soluções para uma determinada classe de problemas, uma evolução da questão da regra tecnológica. Como exemplo, Dresch *et al.* (2015) aponta os conceitos do processo de focalização, ao qual propõe que os sistemas sejam geridos a partir das restrições, a fim de alcançar a meta desejada pela empresa.

Na ciência aplicada o ponto de partida é a criação dos construtos (ou conceitos) que são combinados em construções de ordem mais elevada, gerando modelos que podem ser usados para descrever tarefas, situações ou artefatos. Também são criados métodos como formas de executar atividades orientadas à realização de metas. Em uma instância mais abrangente o design *propositions* se referem então às contribuições teóricas que podem ser feitas por meio da aplicação do design *science research*.

3.2 RELAÇÃO DOS PROCEDIMENTOS À PESQUISA

A partir da metodologia de pesquisa da *Design Science*, são definidos os procedimentos para a pesquisa, conforme Quadro 1.

Quadro 1 – Relação dos procedimentos à pesquisa

ATIVIDADES DSRM	CAPÍTULOS
1. Identificação do Problema e motivação.	Capítulo I: Introdução
2. Definição dos objetivos de uma solução – Inferência.	
3. Projeto e desenvolvimento – Teoria.	Capítulo II: Estado da Arte; e Capítulo III: Metodologia da Pesquisa
4. Demonstração – Conhecimento.	Capítulo IV: Método para a Identificação de Critérios
5. Avaliação – Análise.	
6. Comunicação	Dissertação e artigos

Fonte: dos autores.

O ponto de entrada dessa pesquisa está **centrado no problema**, especificamente na atividade de Avaliação de Ideias no *front end* da Inovação. A partir da definição do problema, são **inferidos** os objetivos da pesquisa, conforme identificado no Capítulo I.

O Capítulo II apresenta o conhecimento necessário para a **teoria**, especialmente para a compreensão dos sistemas de Gestão de Ideias e o processo de avaliação. A partir das revisões sistemáticas, o capítulo permite a identificação de uma oportunidade de desenvolvimento para a pesquisa, direcionada ao conhecimento coletivo. Entre os assuntos discutidos, apresenta a análise de folksonomias para sistemas de conhecimento coletivo.

Em seguida, o Capítulo III define o projeto de desenvolvimento da pesquisa de acordo com os procedimentos abordados pela *Design Science*. Como resultado é gerado o **conhecimento de como fazer**, demonstrado na primeira parte do Capítulo IV.

O Capítulo IV apresenta ainda os artefatos propostos e a demonstração, sendo: um método, por apresentar “*um conjunto de passos para executar uma tarefa*” (Dresch et al., 2015); e um modelo, para a “*representação do espaço de solução*” (Dresch et al., 2015). São realizadas, então, as **análises** para a avaliação dos artefatos, conforme algoritmos e indicações do Capítulo II. Por fim, a pesquisa é

comunicada pelo documento da dissertação e artigos publicados no decorrer do desenvolvimento.

O APÊNDICE E apresenta um esquema completo da dissertação em relação a essa metodologia.

4 APRESENTAÇÃO E ANÁLISE DOS ARTEFATOS

Neste capítulo é apresentada a proposta de um método para representação e uso do conhecimento empregado nas interações sociais livres. Nos tópicos a seguir, são discutidas ainda as características que podem fundamentar a seleção de critérios.

4.1 A ABORDAGEM DO CONHECIMENTO

Tidd e Bessant (2015) relacionam a inovação diretamente à geração de novos conhecimentos ou a novas ligações entre conhecimentos já existentes. Como fase inicial desse processo, o *front end* da inovação apresenta o conhecimento ainda difuso (Koen *et al.*, 2002). Difuso por se apoiar em um processo organizacional interdisciplinar de gestão de ideias. Esse conhecimento ainda difuso evidencia a dificuldade do FEI em selecionar boas ideias.

Por consequência das incertezas e ambiguidades, a etapa não permite um julgamento resoluto. Andrade (2011) caracteriza as incertezas como “*a ausência ou limitação de conhecimento*”. Ao tratar a incerteza, Kempe *et al.* (2012) representam o déficit de informação como um grau de conhecimento.

Cabe ressaltar que a avaliação e seleção são atividades comuns durante todo o processo da inovação e o que diferencia é o avanço das pesquisas (Cooper, 2008). Porém, se há o avanço das pesquisas há também o aumento do custo, o que gera riscos. Os riscos estão associados ao erro de se manter uma ideia que será rejeitada ou não terá sucesso, ou mesmo desperdiçar uma ideia promissora por falta de conhecimento (Brun *et al.*, 2009).

Ceci (2010), ao tratar as diferenças entre dado, informação e conhecimento, baseado em Fialho (2006), destaca que: “*o conhecimento é a completa combinação de informação, dados e relações que levam os indivíduos à tomada de decisão, ao desenvolvimento de novas informações ou conhecimento e à realização de tarefas*”. A isso é evidenciado que o conhecimento depende do contexto e sua função está relacionada à característica de formato e tempo adequados.

A engenharia e a gestão do conhecimento são facilitadoras para mediar à interdisciplinaridade das áreas, considerando os métodos, técnicas e ferramentas de estudo para a identificação, aquisição, formalização, representação, distribuição, implementação, avaliação e manutenção (Egc, 2004).

Dado o que foi apresentado no tópico de análises de folksonomias, é observada uma oportunidade na extração do conhecimento a partir da atribuição de *tag* pelos usuários. Portanto, a contribuição tecnológica dessa dissertação é apresentar um método que permita a definição de critérios de avaliação a partir das folksonomias.

Para tanto, se fundamenta nos conceitos teóricos da Inovação e das Folksonomias. Os conceitos que alinham a fundamentação da Inovação são: o *front end* da inovação, como a fase inicial do processo de inovação; a Inovação Aberta, que se estabelece no cenário econômico como uma tendência para a busca de conhecimento externo a organização; a fase de Gestão de Ideias, na busca por novos conceitos e oportunidades; e, então, os sistemas de Gestão de Ideias, que se destacam entre as soluções de gerenciamento de ideias na Web.

Em seguida, como solução aos problemas percebidos e pontuados nos tópicos anteriores, as Folksonomias, e para o estudo pragmático dessas: a Análise Semântica, como o estudo do conhecimento explicitado por usuários; a *Folksonomy Mining*, que considera a mineração de dados em estruturas de folksonomias; e o estudo dos hipergrafo na representação de tais estruturas. Tais conceitos e os respectivos autores são apresentados a seguir no Quadro 2.

Quadro 2 – Fundamentação teórica.

FUNDAMENTAÇÃO TEÓRICA				
FUNDAMEN- TAÇÃO	PRINCIPAIS AUTORES	CONCEITOS	PRINCIPAIS AUTORES	
ENGENHARIA E GESTÃO DO CONHECIMENTO	Inovação	<i>Front End</i>	Koen <i>et al.</i> (2002)	
		Inovação aberta	Chesbrough (2003)	
		Gestão de Ideias	Kempe <i>et al.</i> (2012)	
		Sistemas de Gestão de Ideias	Westerski <i>et al.</i> (2011)	
	Folksono- mias	Mika (2007)	Sistemas de Conhecimento Coletivo	Gruber (2008)
			Mineração de folksonomias	Hotho (2010)
			Hipergrafos	Lohmann e Díaz (2012)

Fonte: dos autores.

4.2 A INTEGRAÇÃO DO MÉTODO AOS IMS

No contexto de abertura à inovação, os sistemas Web têm apresentado opções bem recebidas pela comunidade. Entre os sistemas abordados na literatura é possível perceber o uso de fóruns, blogs, *microblogging*, *wikis*, redes sociais, Moodle, mensageiros instantâneos, além de sistemas dedicados como Sistema de Gestão de Ideias, Sistemas de resolução de problemas, Marketplace da Inovação, Sistemas de análise da inovação, entre outros. É mais especificamente durante o processo de ideação é que os IMS ganham destaque a fim de sistematizar o processo de gestão de ideias.

Nesses sistemas, a constante interação social desenvolve uma base de participação colaborativa com grande volume de dados. Considerando a abordagem de Kempe *et al.* (2012), a fase divergente criativa pode receber um acúmulo de ideias geradas e/ou recolhidas. Porém, são nas fases de convergência que são apresentados os principais desafios da gestão de ideias, ou seja, no enriquecimento, avaliação e seleção.

Os aspectos de enriquecimento, avaliação e seleção tomam como princípio a abordagem da interação social expressos por: recursos quantitativos como pontuações, “likes”, “estrelas”, importância, números de comentários, métricas que compõem os rankings das ideias (Baez e Convertino, 2012); e, por vezes consideradas técnicas de processamento auxiliado por máquinas como a partir dos comentários (Westerski *et al.*, 2011), conforme observado no tópico 2.3. Nessas análises é observado que não há um aproveitamento ideal dos conhecimentos envolvidos na avaliação pela comunidade, há somente métricas relacionadas à popularidade.

As comunidades representam uma rica fonte dos mais variados conhecimentos. Aproveitar esses diferentes conhecimentos com diversos indivíduos conectados, compreende a dificuldade apresentada por Gruber (2008) sob os aspectos dos sistemas de conhecimento coletivo.

Para isso, a partir do objetivo de “*propor um método de identificação e análise de critérios para avaliação de ideias a partir do conhecimento coletivo em um sistema de gestão de ideias*”, a engenharia do conhecimento evidencia arcabouços que consideram a semântica no tratamento dos dados.

No contexto da Inovação, esse método apresenta uma apropriação dos conhecimentos da comunidade em benefício de uma atividade ainda explorada apenas por métricas de engajamento social e popularidade,

como discutido por Westerski e Iglesias (2011) e Baez e Convertino (2012).

4.3 A CARACTERIZAÇÃO DO MÉTODO

Alternativas para minimizar o déficit de informações nas fases iniciais do processo de inovação são discutidas por Kempe *et al.* (2012). Porém, os autores tomam como objetivo o desenvolvimento de um algoritmo otimizado para calcular uma classificação incompleta. Na pesquisa, os autores não abordam a aquisição dos critérios, partem do princípio que os critérios já estão definidos pela organização, carecendo apenas as informações necessárias para o adequado julgamento. O algoritmo intercede, então, baseado nos critérios possíveis de serem julgados no momento, ordenando parcialmente as ideias mais qualificadas.

Chelvier *et al.* (2011) identificam que os modelos de tomada de decisão usualmente utilizados, não possuem suporte a avaliação multicritérios e multiavaliadores. Já nos modelos de avaliação matriciais, cada avaliação é um fim em si mesmo, eventualmente são destacadas determinadas áreas por pesos, relevando apenas uma classificação por pontuação.

Portanto, em ambos os modos, o conhecimento empregado ao avaliar é expresso apenas em escalas e se perde em valor semântico. A perda de valor semântico se dá pelo fato de, ao registrar apenas uma nota, o usuário não distingui quais elementos foram observados, ou quais critérios considerou ao avaliar um determinado recurso. Fatores que são construídos mentalmente a partir do seu conhecimento e experiência.

Essencialmente, os modelos de avaliação apresentados na revisão da literatura convergem para a recomendação de critérios específicos, por vezes classificados para determinadas áreas, mas, em todos os modelos, um conjunto comum de critérios é distribuído para a avaliação pela comunidade ou por especialistas. Outro modo é utilizar dados essencialmente provenientes da popularidade e engajamento social, um alcance orgânico da proposta.

Por outro lado, ao pontuar uma avaliação livre, cada usuário deve tomar para si critérios de julgamento próprios, naturalmente. Tais critérios são resultantes de sua experiência, conhecimento tácito e percepção da área, sejam objetivos ou subjetivos (Feroli *et al.*, 2010).

Uma forma de destacar esses conhecimentos empregados seria permitir aos usuários apontar os seus próprios critérios que o levaram a conceder tal pontuação para determinada ideia. Assim, o conhecimento explícito em cada julgamento poderia revelar padrões.

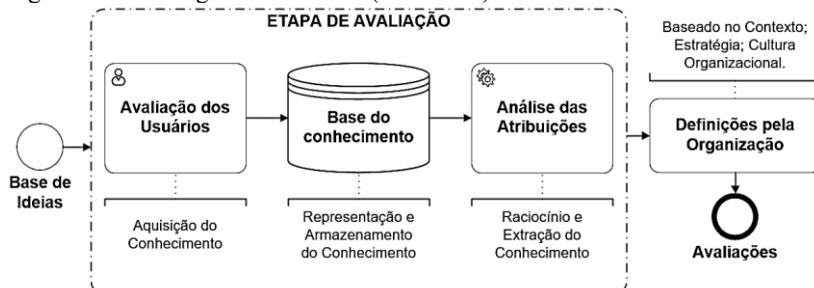
A fim de integrar bases de conhecimento em seus sistemas, Madhusudanan e Chakrabarti (2014), Gómez-Pérez *et al.* (2007), Holder *et al.* (2006) abordam a engenharia do conhecimento como um processo de aquisição, representação e inferência/raciocínio para a solução de problemas. Assim, a partir dessa abordagem da engenharia do conhecimento e tomando a interação social como atividade intensiva em conhecimento, a Figura 17 representa uma visão geral de como se consistiria tal processo.

Nesse esquema, é proposto o início do processo a partir de um conjunto de ideias geradas e armazenadas (Base de Ideias) que despontam de uma Etapa de ideação para uma Etapa de avaliação.

A linha pontilhada destaca a etapa de avaliação como uma marcação de grupo. O grupo delinea uma associação conceptual entre três elementos: (a) a avaliação dos usuários que formam a comunidade online, como uma tarefa de aquisição do conhecimento; (b) a base do conhecimento, como uma etapa de representação e armazenamento do conhecimento; e, por fim, (c) a análise das atribuições, que consiste no raciocínio e extração do conhecimento.

Na primeira ação desse grupo ocorre a aquisição do conhecimento. Usuários manifestam sua avaliação atribuindo o(s) critério(s) considerado(s) e as respectivas notas. As interações desses usuários são representadas por um hipergrafo, apresentado nos próximos tópicos, e registradas em uma base do conhecimento para posterior análise. A análise da base é fundamentada na mineração de dados ou motores de inferência para dar origem a um conjunto de folksonomias.

Figura 17 – Visão geral do Método (Artefato 1).



Fonte: do autor.

Em um momento posterior a avaliação pela comunidade, tais atribuições sociais são então selecionadas por avaliadores internos. Aqui são destacados os fatores que fundamentam a seleção dos critérios, fatores como: uma projeção de contexto; na estratégia; e/ou na cultura organizacional, como destacado por Brem e Voigt (2009).

A saída do processo é então o conjunto de critérios destacados pela comunidade e selecionados pela organização. Nos tópicos a seguir, as ações de cada etapa são mais detalhadas.

4.3.1 Avaliação dos Usuários

A “Avaliação dos Usuários” é identificada no método como a atividade de aquisição de conhecimento. A aquisição do conhecimento consiste na “*criação de novas ideias, reconhecimento de padrões, síntese de disciplinas separadas e no desenvolvimento de novos produtos e processos*” (Cormican e O'sullivan, 2000), sendo considerado um dos grandes desafios da engenharia do conhecimento (Madhusudanan e Chakrabarti, 2014).

Algo semelhante é observado ao avaliar uma ideia. Ao pontuar uma nota, um indivíduo emprega seus próprios conhecimentos em uma construção mental baseada em um conhecimento tácito. Explicitar que critérios são considerados oportunos para avaliação de determinado recurso permite a comunidade, então, colaborar na definição dos critérios de seleção pela organização.

No processo apresentado, a partir da entrada, como um banco de ideias, ocorre a fase inicial do processo, que é caracterizada pela participação da comunidade. Essa participação de um usuário consiste em destacar um ou mais critérios para a avaliação de um recurso, considerado aqui como uma “ideia”.

O critério é definido pela atribuição de uma palavra-chave definida livremente. Além dessa marcação, o usuário concede uma nota à ideia em relação ao critério definido. Como auxílio, é possível recomendar *tags* aos usuários, porém a atribuição deve ser livre e independente, como defendem Mika (2007) e Hotho (2010).

As marcações revelam sinapses que são então armazenadas nas bases de conhecimento. Essas sinapses formarão um conjunto de folksonomias, consideradas como ontologias leves (Mika, 2007), a serem analisadas pelos avaliadores internos.

Cabe distinguir que, as folksonomias apresentam uma dinâmica de participação por um número maior de indivíduos expressando o conhecimento, sintetizado, então, por um conhecimento coletivo. Enquanto uma ontologia formalizada expressa o conhecimento de um número restrito de especialistas.

Assim, segundo Hotho (2010), as folksonomias não sofrem com o afunilamento de aquisição de conhecimento, pois possuem uma significativa disponibilização de conteúdo por muitas pessoas. Essa dinâmica é fundamental, pois permite inferir uma conceptualização compartilhada e atualizar/reciclar a percepção. Afinal, uma ideia se integra ao contexto e o conjunto de fatores acarretados, tais fatores são melhor definidos nos tópicos seguintes.

4.3.2 Base do conhecimento

Após o conhecimento ser adquirido, o desafio é a sua representação. Nessa atividade o conhecimento é estruturado de modo que esteja pronto para ser armazenado, utilizado e compartilhado em uma base (Madhusudanan e Chakrabarti, 2014).

Para a proposta de desenvolver uma análise das folksonomias em IMS, foi necessário compor um modelo que permitisse a emersão de critérios de avaliação. Uma forma de representar a relação entre esses componentes é a partir do conceito de grafo. Por se tratar de uma relação entre diferentes componentes esse grafo é mais bem definido como um hipergrafo (Mika, 2007).

Ao considerar a questão de desempenho, para o armazenamento de dados, os estudos de Bachman (2013) identificam características relacionadas a performance e otimização em banco de dados em grafo.

Bachman (2014) compara uma disposição com propriedades nas arestas, e depois uma organização desses dados em nós. A comparação resulta em performances de busca até duas vezes mais rápidas ao tratar os dados em nós ao invés de propriedades nas arestas.

Portanto, a fim de minimizar o uso de propriedades nas arestas, dentre os modelos apresentados por Lohmann e Díaz (2012), foi escolhida a representação a partir de um hipergrafo tripartido-não-uniforme. Esse modelo minimiza o uso de dados nas arestas e permite a criação de um vértice que converge as informações.

Considerando as representações de marcação discutidas por Lohmann e Díaz (2012), tal hipergrafo foi adequado para o propósito da presente pesquisa. Assim, em um hipergrafo tripartido-não-uniforme a

representação do conhecimento se dá pelo conjunto de instâncias conectadas a um ponto em comum, no caso uma atribuição.

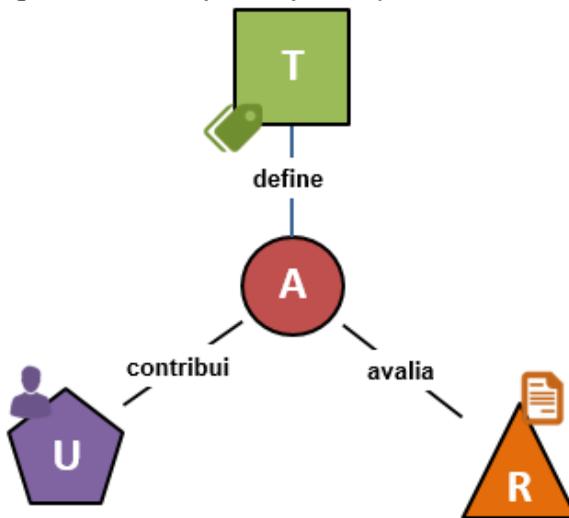
Nessa atribuição foram considerados: o indivíduo como um usuário; o recurso como uma ideia; e a *tag* como um critério. Logo, a representação da atribuição como uma tupla (U, R, T, A) é identificada tal como: U é o conjunto de usuários; R é o conjunto de ideias; T é o conjunto de *tag*; e $A \subseteq U \times R \times T$ é o conjunto de atribuições. Portanto, $a \in A$ é uma tripla (u, r, t) indicando que determinado usuário u atribuiu a *tag* t a uma ideia r , conforme demonstrado na Figura 18.

Também é possível representar a atribuição do modelo com um hipergrafo com a adjacência dada por uma matriz binária tridimensional ao qual $A = \{a_{i,j,k} \mid U \times R \times T\}$ onde cada entrada é dada por $a_{i,j,k} \in \{0, 1\}$. Essa atribuição pode ser então representada como a relação do conjunto.

A partir da representação, a seleção de ideias é enriquecida pela comunidade a partir da interação qualitativa e temporal. Para isso, o nó “Atribuição” considera as propriedades “valor de avaliação” e “tempo”.

Assim, a propriedade “valor” tem a função de qualificar tal atribuição. Portanto, além de atribuir um critério, o usuário considera uma nota na marcação que se integrará ao momento estratégico organizacional. O conjunto das notas é considerado mediante os critérios destacados nas etapas seguintes.

Figura 18 – Modelo para a representação do conhecimento (Artefato 2).



Fonte: do autor.

Outro dado inserido no nó “Atribuição” é o “tempo”. A importância do registro do tempo para análise de folksonomias é amplamente discutida desde Dubinko *et al.* (2007). O intuito de registrar essa propriedade é de permitir uma reconstrução da linha temporal para os eventos de marcação, tal que seja possível analisar a estabilização da folksonomia de critérios de avaliação e relevância destacada pelos usuários nas atribuições.

Para a marcação temporal, o padrão *timestamp* do UNIX é identificado como alternativa aos algoritmos. O tempo Unix, ou “era POSIX”, descreve instantes de tempo definidos como o número de segundos desde as 00:00:00 do dia 01º de janeiro de 1970 na linha do Tempo Universal Coordenado (UTC).

A propriedade também é utilizada por Jelassi (2012) a fim de identificar tendências e descobrir a relevância dos usuários, *tag* e recursos em determinados períodos. O autor propõe então um algoritmo denominado QUADRICONS que considera o tempo como uma quarta dimensão.

Mika (2007) cita, que, na prática, folksonomias são muito menos voláteis do que seria de se esperar, o autor cita que isso é devido "*a maioria dos termos ter um significado único bem definido que se mantém estável durante longos períodos de tempo*". Um exemplo dado é o site *del.icio.us*, ao qual muda diariamente a sua base, porém a comunidade como um todo se altera muito mais lentamente.

Portanto, a partir da interação social no tempo, as folksonomias criam uma "*espinha dorsal estável e substancial*" (Mika, 2007). Para o autor, essas partes estáveis podem ser consideradas como "*ontologias leves que representam um consenso sobre o significado dos termos*", pois se assemelham as ontologias que se estabilizam após longo debates.

As propriedades e as interações sociais convergem para a base do conhecimento, onde são armazenadas. O foco dessa pesquisa não é definir o tipo de base a ser utilizada, contudo, pela representatividade da análise em grafo, bancos de dados em grafos são recomendações a serem refletidas pelo analista.

O armazenamento em uma base de conhecimento permite a exploração a fim de se obter um conhecimento aplicado. Pois, considerando as redes formadas nas relações das folksonomias e representadas por um hipergrafo tripartido, é possível analisar a semântica das marcações e as interações sociais por intermédio das informações no processo de análise das atribuições.

4.3.3 Análise das atribuições

A partir do conhecimento adquirido e representado, esse deve servir às finalidades dos sistemas, tais como na realização de uma análise e extração do conhecimento. Isso é feito por meio dos processos de inferência, ou raciocínio (Madhusudanan e Chakrabarti, 2014). Para Madhusudanan e Chakrabarti (2014), soluções baseadas em conhecimento necessitam de dois componentes: (a) uma base de conhecimento, que consiste essencialmente do conhecimento especializado estruturado; e (b) um motor de inferência, para processar e realizar o raciocínio pretendido.

Segundo Tanwar *et al.* (2010): "*redes semânticas são uma representação gráfica declarativa que podem ser usadas tanto para representar o conhecimento ou para apoiar sistemas automatizados para raciocinar sobre conhecimento*".

Ao considerar essas redes em uma base consolidada, há a necessidade de tratar e extrair o conhecimento aplicado. Para isso, algoritmos auxiliam na inferência, como: redes bayesianas, herança, classificação, sistemas de recomendações, entre outros. Uma abordagem aprofundada dessas inferências baseadas em grafos conceituais é descrita já por Delugach e Hinke (1994).

Outros autores como Holder *et al.* (2006), Geldart e Cummins (2010), Tsui *et al.* (2010), Jelassi (2012) e Bastian *et al.* (2014) apresentam análises a partir de modelos em grafos para a inferência de conhecimento e raciocínio baseados na adaptação de algoritmos e técnicas.

Holder *et al.* (2006) aborda a inferência, em especial com técnicas para determinar o valor de incerteza associados a lógica. Para isso, além das redes Bayesianas cita como abordagem emergente e promissora as redes lógicas de Markov (Richardson e Domingos, 2006).

Geldart e Cummins (2010) discutem um método que permite a inferências sobre folksonomias utilizando a lógica não axiomática. Essa técnica trata-se de uma semântica baseada em evidências que opera sob a hipótese de conhecimento e recursos insuficientes. O resultado demonstra que a técnica se mostra promissora para a integração de sistemas de *tag* com representações de conhecimento de modo mais estruturado.

Bastian *et al.* (2014) apresenta uma solução para inferência e recomendação de competências a partir das folksonomias de habilidades, as folksonomias são extraídas das atribuições de perfis na

rede social LinkedIn®. Nessa pesquisa, os autores identificam uma padronização dos atributos nos perfis tais como a companhia, títulos, indústria, área de estudo ou filiação a grupo, entre outras informações úteis na inferência de habilidades ou áreas de expertise utilizando a técnica de redes Bayesianas em folksonomias.

Em Tsui *et al.* (2010) é proposto um método para converter automaticamente *tag* em uma taxonomia hierárquica. O método proposto envolve uma análise heurística e regras de um algoritmo de aquisição de conceito-relação. Nessa abordagem há a inferência entre os padrões extraídos de modo a deduzir relações ocultas entre as *tag* que não estão diretamente descritas nos recursos. O resultado mostrou um aumento nos parâmetros de revocação e precisão em relação a métodos de construção de taxonomia comum.

Jelassi (2012) extrai das folksonomias o tempo, considerando como uma quarta dimensão, útil para derivar regras quadráticas, raciocinando sobre essa dimensão, a fim de perceber as relações de saída entre as *tags*, recursos e usuários.

Em comum entre os autores, o propósito é permitir que folksonomias expressem resultados automatizados de padrões, relações ou recomendações.

Segundo Hotho (2010), as folksonomias não explicitam determinadas conceptualizações compartilhadas, mas também não forçam os usuários a utilizar *tags* consistentes. Apesar disso, o uso de *tags* dos usuários com interesses semelhantes, tende a convergir para um vocabulário compartilhado.

Como apresentado, o resultado da interação social livre é uma folksonomia representada por um hipergrafo que pode ser decomposto para diversas análises com base no estudo de análise de grafo discutidas no tópico 2.5. Assim, o modelo busca descobrir conceituações compartilhadas que emergem da marcação social e que estão ocultas em uma folksonomia.

Essa análise resulta em um conjunto de critérios de avaliação baseado nas métricas de análise em grafo. As métricas direcionam a análise das avaliações recebidas a partir da comunidade. O conjunto de dados gerados pelas atribuições permite, assim, análises exploratórias dos dados de forma qualitativa e quantitativa, para, a partir de anotações prévias das atribuições: sugerir critérios, recomendar usuários, qualificar, ou analisar tendências.

Para obter uma melhor compreensão do conhecimento extraído, destas relações marcadas entre *tags*, é proveitoso que usuários e recursos sejam ligados a partir de um mapeamento por fontes externas de

conhecimento. Estas fontes providenciam então o significado semântico mais claro ou o relacionamento melhor fundamentado do conhecimento representado (Hotho, 2010).

As três dimensões propiciam o uso de tais fontes externas, o que fornece uma semântica das informações capturadas por ontologias de dados léxicos e tesouros, por exemplo. Estas informações são então abordadas para a avaliação do significado de diferentes medidas de similaridades.

Assim, as propriedades do grafo, em sistemas de gestão de ideias, ampliam as oportunidades de avaliação e seleção de ideias em todo o processo de inovação. Pois, permite uma análise relacionada aos motivos das atribuições qualitativas e uma definição pela organização de quais critérios satisfaz o seu contexto.

4.3.4 Definições pela Organização

A etapa de “Definições pela Organização” parte da análise das atribuições. O princípio é considerar o contexto organizacional como seleção dos critérios que efetivamente se constituem em uma preocupação para a escolha das ideias, como utilizado por Brem e Voigt (2009).

Segundo Madhusudanan e Chakrabarti (2014) o "*contexto denota a aplicabilidade do conhecimento, ou seja, para decidir a respeito de que pedaços de conhecimento são válidos sob um dado conjunto de circunstâncias*". O autor destaca que tanto o "contexto" como o "conhecimento" podem ser explícitos ou implícitos. Sendo que, enquanto o conhecimento permanece relativamente inalterado entre as fases de aquisição e utilização do conhecimento, o contexto acionado por um determinado acontecimento, se altera durante a ação.

Madhusudanan e Chakrabarti (2014) descreve que o “contexto” pode ser visto como uma forma de conjunto de conhecimento para: (a) busca eficiente; (b) representação de situações contrafatuais ou hipotéticas; (c) compreensão dos efeitos de ações específicas para situações particulares; e (d) dirigir o foco de atenção de um agente aos aspectos mais salientes de uma situação. Ainda segundo o autor uma situação “*representa o contexto dentro do qual o conhecimento adquirido de um especialista é válido*”.

Em um pensamento sistêmico, Hester e Adams (2014) consideram o contexto em cinco elementos essenciais: (1) circunstâncias; (2) fatores; (3) condições; (4) valores; e (5) padrões.

Dentro desses elementos podemos destacar, por exemplo, o cenário atual, as estratégias da empresa e a cultura organizacional.

Em relação ao método apresentado, ao sistematizar o conhecimento dos critérios de avaliação, o intuito é agregar a construção mental dos usuários em um processo de tomada de decisões. O que dinamiza uma análise semântica de acordo com os objetivos da organização. Assim, ao analisar uma avaliação, um conjunto de critérios apontados pelos usuários é apresentado por uma folksonomia.

Tal folksonomia de critérios é ainda analisada segundo os objetivos da empresa, como auxílio à tomada de decisão. A partir dessa associação é possível criar inferências e analisar aspectos que auxiliem nos processos da inovação.

Portanto, permite ao gestor selecionar as ideias não apenas pela popularidade ou pontuações conquistadas, mas tomar as decisões de acordo com o contexto descrito. Assim, a tarefa é analisar as folksonomias de critérios e definir quais são relevantes para a avaliação e seleção das ideias marcadas.

Mesmo não havendo condições para avaliar os determinados critérios logo nas fases iniciais, a continuidade da ideia permite direcionar as pesquisas em relação aos critérios relevantes já marcados pela comunidade. A partir dos avanços nas demais fases, as incertezas podem ser reduzidas por um direcionamento das pesquisas, já que certos critérios são imensuráveis nas fases iniciais, porém essenciais no decorrer do processo de inovação. Assim, considera Ferioli *et al.* (2010) sobre a distinção de determinados critérios, ao qual são julgados objetivamente ou subjetivamente.

No contexto econômico, Andrade (2011) define que "*o grau de confiança está diretamente associado ao grau percebido de incerteza do conhecimento relacionado a um possível evento futuro*".

Ao destacar "evento futuro", é possível considerar o mesmo diante dos cenários e as incertezas do processo, como descrito por Sandström e Björk (2010) que também evidenciam que existe um risco de que ideias sejam filtradas não por serem consideradas ruins, mas porque "*não se encaixam no modelo de negócios atual e processo de avaliação*".

Portanto, a proposta da base de conhecimento não é descartar ideias, mas enriquecê-las para que tenham valor no momento adequado. E, para isso, são necessárias técnicas de tomadas de decisão sobre o contexto, em especial aquelas que abordam as incertezas dos dados.

Chelvier *et al.* (2011) identificam que, em geral, os dois métodos que podem ser utilizados nas fases iniciais de um processo de inovação

são: o *Analytic Hierarchy Process* (AHP) (Saaty, 1988) e *Cost-Benefit Analysis* (CBA) (Chakravarty, 1989). Essas técnicas são vistas como uma continuação da presente pesquisa, voltada especificamente para a seleção de ideias.

4.4 A VERIFICAÇÃO DO MÉTODO PROPOSTO

O estado da arte apresentou alguns trabalhos mais recentes em relação a mineração de dados sob as interações sociais, técnicas apresentadas como “*Folksonomy Mining*”. Na pesquisa é destacado o uso de algoritmos e estatísticas em especial para: (a) Decomposição de grafos; (b) Classificação; (c) Distribuição; e (d) Recomendação.

A partir do modelo apresentado, tais análises são verificadas observando a aplicação para o domínio de avaliação de ideias. Como etapa de verificação, esse tópico considera a representação da interação social em hipergrafo. A representação é utilizada por diferentes ferramentas de análise.

Nos tópicos a seguir são discutidos: as observações, o conjunto de dados utilizados, a adequação ao modelo, as análises e os resultados.

4.4.1 O Conjunto de Dados

O conjunto de dados adotado para a análise foi disponibilizado no portal de arquivos do GroupLens, laboratório de pesquisa no Departamento de Ciência da Computação e Engenharia da Universidade de Minnesota. Dentre os arquivos liberados, o mais adequado para a verificação da pesquisa foi o conjunto de dados intitulado "MovieLens + IMDb/Rotten Tomatoes", disponibilizado por Cantador *et al.* (2011).

Devido à ausência de um Sistema de Gestão de Ideias com a abordagem das folksonomias, o conjunto de recursos como filmes, ao invés de ideias, foi o mais adequado para a realização das análises abordadas pela literatura. Esse uso não interfere nas análises das atribuições, senão pela dinâmica diferenciada do tipo de recurso analisado. Consequentemente, permite observar a generalização do método proposto em diferentes aplicações.

A análise com esses dados tem por objetivo verificar o modelo e observar o comportamento em grandes conjuntos de atribuições a partir dos algoritmos e técnicas apresentados na literatura.

As informações de Cantador *et al.* (2011) foram extraídas das bases de serviços online de avaliações de filmes e traz também as

atribuições de etiquetas de texto. No total são informações de 2.113 usuários, 10.197 filmes, 855.598 avaliações, 13.222 *tags* e 47.957 atribuições, extraídos de serviços online de avaliação de filmes.

O conjunto de dados de Cantador *et al.* (2011) é disponibilizado em um pacote de arquivos compactados. Após a extração, cada arquivo possui a extensão “*.dat”, que é uma extensão genérica para arquivos binários. Dos 12 arquivos do pacote compactado, sete possuem informações de filmes extraídas dos sites MovieLens, IMDb e Rotten Tomatoes, outros quatro trazem informações das interações sociais e um último arquivo apresenta as informações das *tags*/palavras chaves utilizadas pelos usuários.

O Quadro 3 apresenta os arquivos "movies.dat", "tag.dat", "user_taggedmovies-timestamps.dat" e "user_ratedmovies-timestamps", utilizados para compor a análise, os demais arquivos foram omitidos.

Quadro 3 – Arquivo e propriedades do conjunto de dados.

ARQUIVO	PROPRIEDADES
tag.dat	id; value
movies.dat	id; title; imdbID; entre outras.
user_taggedmovies-timestamps.dat	userID; movieID; tagID; timestamp
user_ratedmovies-timestamps.dat	userID; movieID; rating; timestamp

Fonte: do autor, adaptado de Cantador *et al.* (2011).

No arquivo “tag.dat”, o autor apresenta os dados da classe “tag” com a indicação das palavras-chave que constam no dataset. Semelhantemente, o arquivo “movies.dat” apresenta os dados da classe “movie” com as informações das instâncias de filmes que constam nas plataformas.

No arquivo “user_taggedmovies-timestamps.dat” constam as informações das *tags* atribuídas pelos usuários a um determinado filme. Nos serviços, tais *tags* são utilizadas para categorizar o portfólio de informações dos filmes e auxiliam o usuário a indexar seu portfólio, semelhante aos demais serviços de atribuição social como o *del.icio.us*.

Seguindo o modelo de avaliação dos filmes por atribuição de nota, o arquivo “user_ratedmovies-timestamps.dat” traz as informações em relação a pontuação atribuída.

Os dados foram então organizados em uma planilha Excel. A associação gerou um conjunto de dados adaptados para ao modelo apresentado no tópico 4.3.2. Essencialmente as alterações ocorreram nas propriedades, reduzindo informações, agregando os dados, porém mantendo as relações.

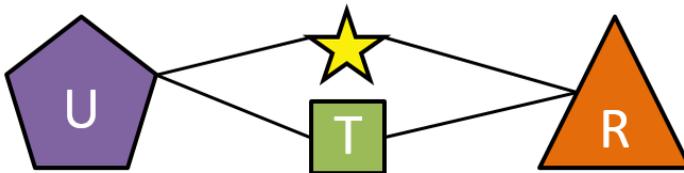
Na classe “atribuição”, principal objeto da análise, foi necessária uma adaptação. Os serviços utilizados para extrair os dados consideram que: um usuário avalia com uma nota; e pode atribuir uma etiqueta para categorizar ou indexar.

Portanto: “O usuário ‘U’ atribuiu a um filme ‘R’ a pontuação ‘X’.”, dados constam no arquivo “*user_ratedmovies-timestamp.dat*”; e “O usuário ‘U’ atribuiu a um filme ‘R’ a tag ‘T’.”, dados constam no arquivo “*user_taggedmovies-timestamp.dat*”), conforme Figura 19 onde: “U”, representa o usuário; “T”, a tag; e “R”, o recurso, no caso um filme.

O método proposto considera que a etiqueta é um critério observado pelo usuário para avaliar o recurso, e, portanto, etiqueta e nota são associados. O ideal proposto pelo método de atribuição colaborativa é que ao atribuir uma tag o usuário pontue uma nota, portanto: “O usuário ‘U’ atribuiu a um recurso ‘R’ a pontuação X pela tag ‘T’.”. Assim, a etiqueta é um critério a ser considerado e o conjunto das atribuições formam as folksonomias. A Figura 20 apresenta os dados adaptados ao modelo proposto.

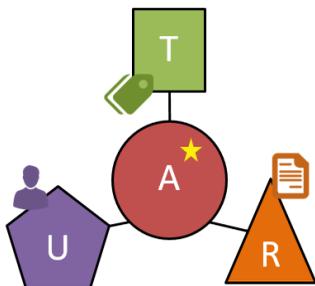
Não necessariamente um usuário que pontuou um filme marca uma tag, assim como não obrigatoriamente o usuário marca uma nota a um filme que atribui uma tag. Então, a fim de completar o esquema necessário para a análise, as atribuições consideraram apenas marcações em que algum usuário tenha pontuado e também atribuído uma tag a algum filme. Foram então atribuídas para as etiquetas, as mesmas notas que o usuário avaliou o filme.

Figura 19 – Atribuições de nota e etiquetas desassociadas, conforme conjunto de dados.



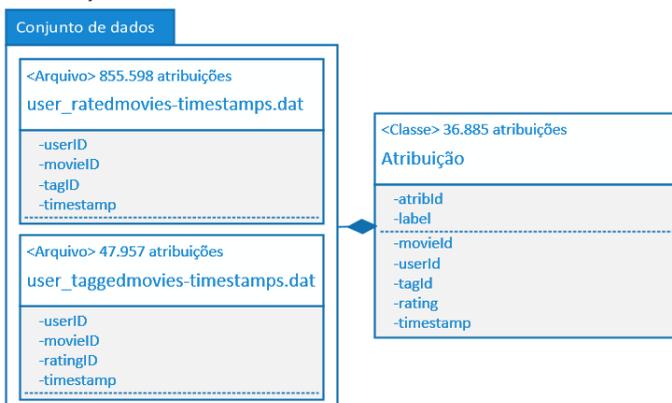
Fonte: do autor.

Figura 20 – Atribuições de avaliação e etiquetas associadas, conforme o modelo proposto.



Fonte: do autor.

Figura 21 – Junção das propriedades de diferentes arquivos na classe “Atribuição”.



Fonte: do autor.

Ou seja, foram mantidas apenas as contribuições ao qual um mesmo usuário atribuiu a nota para o filme e também atribuiu alguma etiqueta para o mesmo filme. Essa exclusão foi necessária, para adequar ao procedimento apresentado nessa pesquisa, ao qual o usuário atribui junto com a etiqueta uma pontuação.

Assim, as propriedades utilizadas para os nós de atribuição foram: “id”, uma identificação única ordenada com base nas atribuições mantidas; “class”, identificados como “atrib”; “userID”, identificação do usuário que atribuiu; “movieID”, recurso associado; “tagID”, etiqueta marcada; “timestamp”, marca do tempo; e “Rating”, a pontuação em uma escala de 0,5 a 5,0, conforme Figura 21.

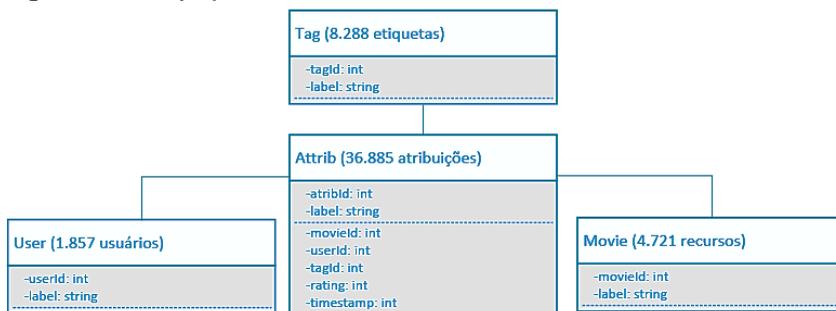
Devido a essa associação, mantiveram-se 36.885 atribuições de etiquetas, associadas com avaliações pontuadas pelos respectivos usuários. Participaram dessas atribuições: 1.857 usuários, 4.721 recursos/filmes e 8.288 etiquetas. As demais instâncias foram omitidas para redução da análise.

Outra adaptação do conjunto de atribuições foi em relação a propriedade de tempo. No conjunto de dados disponibilizado por Cantador *et al.* (2011), o *timestamp* representa a interação temporal dos usuários na marcação de *tag* para filmes com precisão de segundos. Como tal precisão não apresenta variação significativa para a análise, a propriedade foi reduzida para cinco dígitos. Para essa conversão foi calculada na planilha com: $((\text{'VALOR'}/60)/60)/24 + \text{DATE}(1970,1,1)$. Devido a essa conversão, foram ignorados os dados de: hora, minuto e segundo; e mantidos: dia, mês e ano.

Das classes “*tag*”, “*movie*” foram reduzidas para: “*id*”, a partir da chave primária correspondente no dataset; “*rotulo*”, ao qual foi acrescentado uma cadeia de caracteres com a letra “*T*” e “*M*”, respectivamente, concatenada a um número sequencial, exemplo: “*T12*” e “*M34*”; “*class*”, sendo “*movie*” ou “*tag*”; e “*valor*”, que manteve os dados de palavra-chave (“*value*”) e título em inglês (“*title*”), respectivamente.

Por serem omitidos do dataset os dados dos usuários, para a classe “*user*” foram agregadas apenas as propriedades encontradas no arquivo “*user_taggedmovies-timestamps.dat*”. São consideradas então: “*id*”; e “*rótulo*”, ao qual foi acrescentado um número sequencial com a concatenação da cadeia de valor “*U*”+(um número sequencial), exemplo: “*U513*”. Portanto, o resultado obtido da adequação é apresentado na Figura 22.

Figura 22 – Adequação dos dados.



Fonte: do autor.

4.4.2 A análise das folksonomias

Para a análise das folksonomias, os arquivos do conjunto de dados foram carregados: (a) em planilha eletrônica Microsoft Office Excel®; (b) no sistema de banco de dados em grafo Neo4j; e, também, (c) no software Gephi.

No sistema Neo4j, as ferramentas de análise não estão incluídas, sendo necessário adicionar algoritmos, porém, se mostra pela literatura, uma opção adequada em banco de dados para o armazenamento, em especial dados que geram grafos e redes sociais (Bachman, 2013).

Dentre esses artigos, uma alternativa indicada para a análise de dados do Neo4j é a integração com o Gephi. O módulo “*Neo4J Graph Database Support*” apresenta uma opção para isso, permitindo a importação, filtragem e exportação dos dados.

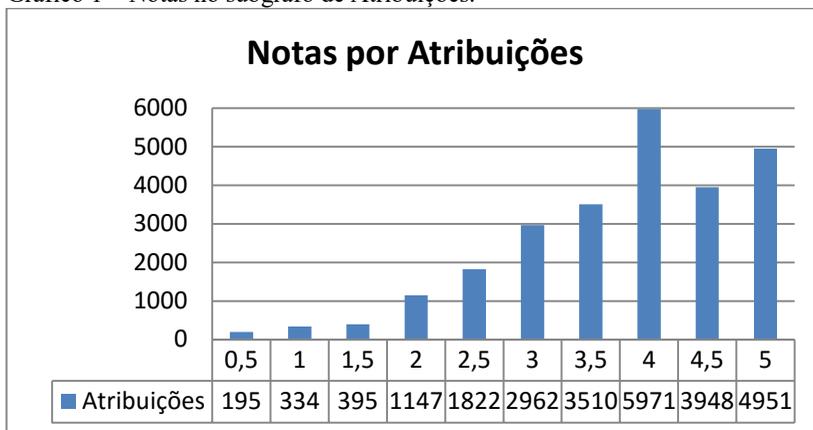
No software Gephi, por ser originalmente uma ferramenta para a análise de grafos, há algoritmos, ferramentas, estatísticas, filtros, módulos e recursos visuais que permitem a observação. A partir da aba “Laboratório de dados”, os grafos resultantes podem ser exportados para planilha eletrônica. Em planilha, a análise permite destacar os resultados com os nós e propriedades em forma de tabela.

Utilizando esses recursos, é possível realizar os procedimentos para obter: (a) Decomposição de grafos; (b) Classificação; (c) Distribuição; e (d) Recomendação. Em cada uma dessas técnicas foram utilizados ao menos um dos algoritmos apresentados para tal fim, devido à limitação da presente pesquisa em se aprofundar em cada recurso. Além disso, a partir do objetivo de analisar a interação social a partir do modelo proposto, os rótulos e títulos foram omitidos.

Antes de iniciar essas análises, o grafo gerado pelo conjunto de dados foi filtrado com o algoritmo *k-core*. O *k-core* gera um subgrafo ao qual cada nó está associado a pelo menos “*k*” vértices no subgrafo (Sariyüce *et al.*, 2013). O algoritmo é comumente utilizado para análise de redes complexas, tais como na detecção de comunidades. Para redução do grafo foi adotado $K=3$.

O algoritmo reduziu o grafo para 30.067 vértices ligados por 75.705 arestas, 82% dos dados carregados. A pequena redução se deve ao fato da alta concentração que apresentam os nós e por representar um único domínio.

Gráfico 1 – Notas no subgrafo de Atribuições.



Fonte: do autor.

a) **Decomposição:**

Com esses dados, a primeira análise destacada é a decomposição dos hipergrafos em subgrafos. A decomposição é fundamental para a observação mais detalhada das instâncias. Na ferramenta Gephi, essa decomposição pode ser realizada a partir de filtros. O filtro “Partição” na biblioteca “Atributos” separa os nós a partir da propriedade.

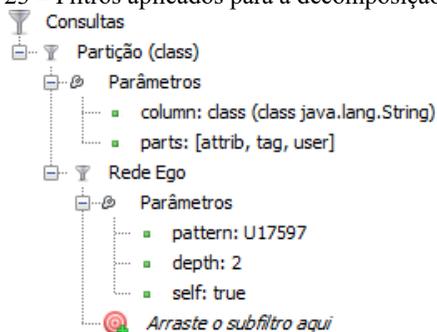
A primeira análise se atém na pontuação atribuída pelos usuários, com o propósito de conhecer o comportamento das avaliações, ou seja, qual as notas concedidas ao atribuir uma avaliação. A análise é tomada a partir do subgrafo das atribuições (class: “attrib”). Como elemento central da análise, uma das propriedades destacada é a média ponderada das atribuições: 3,8 de uma escala de 0,5 a 5 pontos. O Gráfico 1 apresenta a quantidade de atribuições por pontuação.

A análise do Gráfico 1 permite observar que o maior número de participações está concentrado em avaliar os recursos e as *tags* entre quatro a cinco pontos (59%). Portanto, para efeito de análise, pontuações dentro desse intervalo serão consideradas positivas, enquanto pontuações inferiores serão consideradas negativas.

Outra análise fundamental a partir da decomposição é o subgrafo das atribuições e *tags* (AT). Essa decomposição identifica o comportamento de um ou mais usuários ao utilizar determinada tag em sua personomia. Para realizar essa análise, a partir do conjunto de dados, é composto o subgrafo de palavras-chaves de um determinado usuário, escolhido aleatoriamente. O subgrafo AT do usuário U17597 é gerado

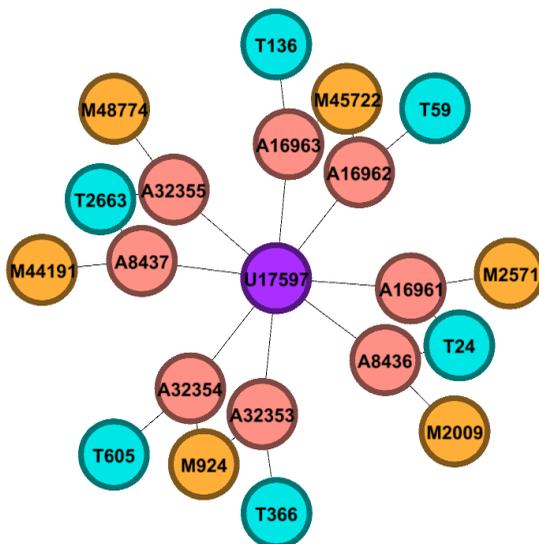
com filtros, considerando: a Rede Ego, com a identificação do nó e definida a profundidade para dois níveis; e, em uma camada superior, adicionado o filtro para a Partição de entidades “user”, “attrib” e “tag”, conforme Figura 23. O subgrafo resultante é observado na Figura 24.

Figura 23 – Filtros aplicados para a decomposição do grafo em subgrafo.



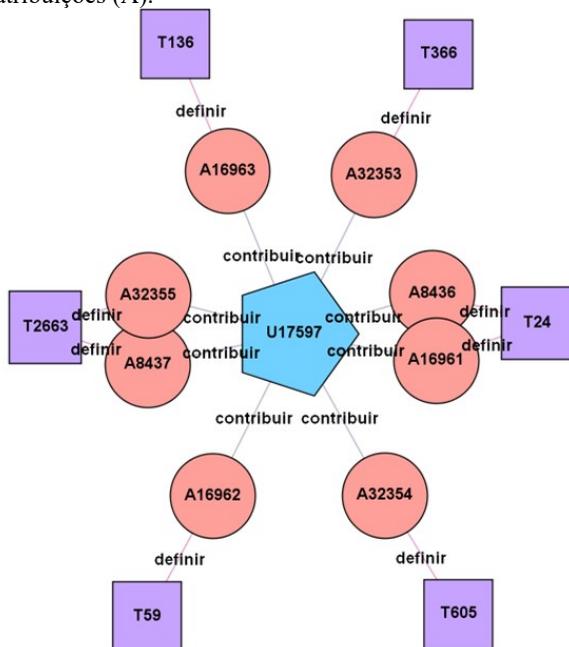
Fonte: Captura de tela do Gephi.

Figura 24 – Subgrafo do usuário selecionado.



Fonte: Captura de tela do Gephi.

Figura 25 – Subgrafo AT do usuário (U) com as etiquetas (T) utilizadas nas atribuições (A).



Fonte: do autor.

Ao desconsiderar os recursos na análise, a Figura 25 demonstra o subgrafo do usuário U17597 com as suas atribuições (A) associadas com as etiquetas (T), a visualização do subgrafo AT é analisada com o algoritmo de Hu (2005).

Assim, é determinada a rede de filiação de um usuário específico e o seu comportamento nas marcações. A rede de filiação é uma das técnicas discutidas por Mika (2007). O subgrafo apresenta ainda o número de vezes que uma determinada *tag* foi atribuída pelo usuário.

Uma forma de redução proposta para a análise seria determinar as coocorrências como um peso nas arestas, excluindo as informações das atribuições (Tabela 1), formando o subgrafo UT. Nesse caso, cada uma das ligações com as etiquetas T2663 e T24 teria peso dois.

Tabela 1 – Rede de filiação do usuário.

UT	T24	T59	T136	T366	T605	T2663
U17597	2	1	1	1	1	2

Fonte: do autor.

No subgrafo UT, a rede de afiliação, apresenta uma transformação ponderada pelo número de instâncias. O grau de cada *tag* é utilizado como peso das arestas, conforme apresentado na Figura 26.

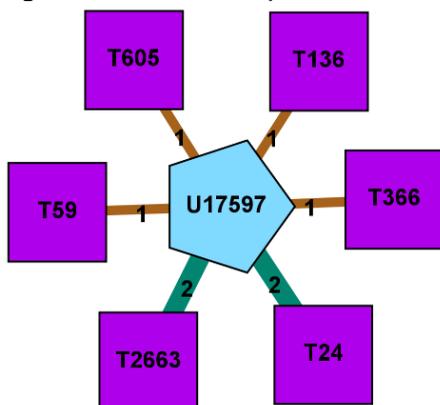
b) Classificação:

A partir da decomposição, é possível analisar mais especificamente os nós com as classificações. As classificações podem representar as convergências que ocorrem a partir das folksonomias e, para isso, diversos algoritmos e estatísticas da rede podem ser utilizados, conforme discutido no capítulo 2.

Considerando o modelo proposto, a busca de uma tríade pode aplicar o filtro da “Rede Ego” com a profundidade para dois níveis, ao qual permite navegar entre as instâncias de atribuição (“attrib”), até o próximo recurso associado. Com essa técnica é possível coletar as informações dos subgrafos gerados, bem como as ideias e usuários conectados.

No caso analisado, ao particionar para o subgrafo “Atribuição/Tag” (AT), foram 4,5 atribuições por etiqueta. Filtrando as etiquetas para aquelas que tenham três ou mais ocorrências, são 2.238 etiquetas que somam 29.567 atribuições, com média de 13,2 atribuições por etiqueta. Já, ao desconsiderar usuários e recursos com menos de três participações, há uma relação “Atribuição/Recurso” média de 7,8 atribuições por recurso. Dentre essas atribuições, as etiquetas de texto com maior grau (*degree*), ou seja, que houveram um maior número de atribuições, possuem um grau superior a 37 contribuições.

Figura 26 – Rede de afiliação do usuário U17597.



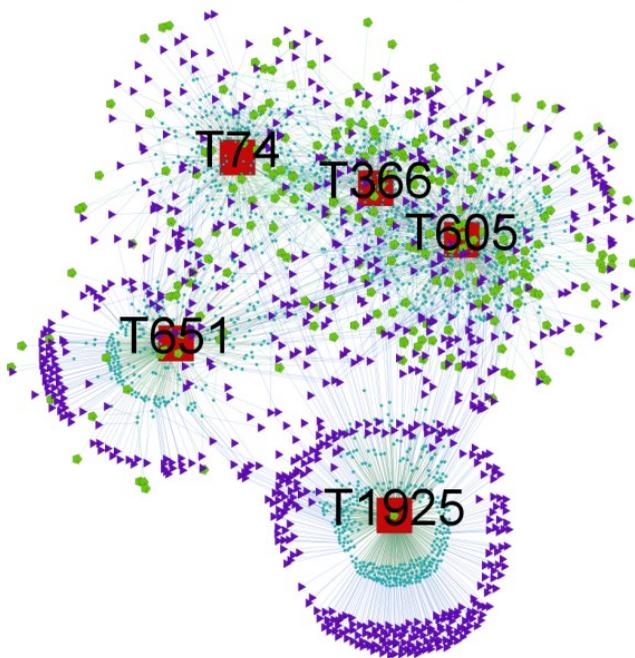
Fonte: do autor.

Tais números refletem a fragmentação das *tags*, afinal, as contribuições pela comunidade são livres e, portanto, cada usuário atribui as suas observações de acordo com a sua personomia. Apesar disso, é possível perceber conjuntos de *tags* que são mais utilizadas, ou mesmo comunidades que convergem em combinações semelhantes. Tal observação é essencial para analisar as características utilizadas por grupos de usuários, identificando o público alvo.

A partir da filtragem das cinco etiquetas com maior número de atribuições, podemos observar esse alto grau de associação. O que representa o desenvolvimento de uma folksonomia com certa convergência entre os usuários, por atribuir etiquetas semelhantes para determinados recursos, conforme Figura 27, onde polígonos de cinco pontos indicam os usuários e de três pontos indicam os recursos.

Ao considerar o conjunto completo, a Tabela 2 apresenta a quantidade de atribuições na escala de notas arredondada que cada *tag* recebeu. Na tabela é possível observar ainda a média geral e o total de contribuições que marcaram a *tag*.

Figura 27 – Grafo das cinco *tags* com maior grau da rede.



Fonte: do autor.

Tabela 2 – Dados das contribuições associadas as 10 *tags* mais atribuídas, ordenadas pelo grau do nó na rede.

<i>Tag</i>	NOTAS NAS ATRIBUIÇÕES					GERAL	
	Rótulo	1	2	3	4	5	Média
T605	-	3	23	122	263	4,6	411
T1925	1	2	55	197	102	4,1	357
T651	-	4	21	100	104	4,3	229
T74	4	14	32	51	109	4,2	210
T366	1	6	29	58	104	4,3	198
T583	2	5	7	62	119	4,5	195
T6800	2	9	32	104	44	3,9	191
T1023	9	16	46	64	34	3,6	169
T247	-	12	21	66	66	4,1	165
T712	1	2	14	54	90	4,4	161

Fonte: do autor.

A análise das médias permite identificar *tags* utilizadas para pontuar um recurso como positivamente ou negativamente pelo grupo, acima ou abaixo da média, respectivamente. Por exemplo, a *tag* mais utilizada, T605, recebeu média de 4,6 pontos, superior à média geral que é de 3,8 pontos. Portanto, a *tag* T605 pode ser definida como essencialmente positiva para o domínio. Já a T1023 recebeu média de 3,6 pontos, ou seja, abaixo da média geral, o que pode ser considerado como um critério pontuado essencialmente como negativo para o domínio.

As etiquetas com maior número de atribuições são também as mais positivas, sendo, em ordem: T605, T1925, T651, T583 e T712.

Além dessa análise, é possível perceber quais etiquetas são mais atribuídas para avaliações abaixo da média. A Tabela 3 apresenta as *tags* ordenadas pelo número de atribuições inferior a quatro.

O grande número de atribuições, com baixas avaliações, aponta que tais etiquetas estão associadas a critérios considerados negativos para o domínio. Assim, sua ocorrência pode representar um ponto fraco a ser considerado para melhorias, por exemplo.

No caso das *tags*, a importância que tomam em relação ao grafo revela a necessidade de serem abordadas como definição de critérios em formulários de avaliação. Tais *tags* são destacadas pela comunidade nas mais importantes atribuições, portanto apresentam aspectos fundamentais para o conjunto.

Tabela 3 – Etiquetas pontuadas essencialmente como negativas.

<i>Tag</i>	NOTA	CONTRIBUIÇÕES			
Rótulo	Média	< 4	> 4	Total	Diferença
T369	1,8	115	4	119	-111
T662	3,1	85	27	112	-58
T2538	3,3	101	54	155	-47
T195	1,9	50	4	54	-46
T2261	3,0	45	6	51	-39
T93	3,3	96	57	153	-39

Fonte: do autor.

Porém, o alto grau das etiquetas não representa fielmente a importância para a rede. Uma estatística alternativa seria então o algoritmo FolkRank. Nessa abordagem há uma variação do algoritmo PageRank, como destacado por Hotho (2010).

A partir das estatísticas do PageRank é possível conferir a importância do nó na rede, pois, um alto grau de associação não reflete confiavelmente a importância da atribuição. Como exemplo, a partir do subgrafo apresentado na Figura 27, é possível perceber que a *tag* T1925 apresenta um alto grau na rede, porém poucos usuários associados. Logo, representa um termo específico ao grupo e não uma importância para o conjunto.

A fim de evitar que isso ocorra, ao considerar o nó intermediário – “atrib” – a associação de um usuário, etiqueta e recurso dividem a importância, semelhante ao utilizado pelo algoritmo PageRank. Desse modo, a importância do nó intermediário é dividida entre os nós associados.

Por outro lado, se não houvesse um nó intermediário, um usuário pode ter a classificação inflada por simplesmente utilizar uma *tag* considerada importante na rede. O mesmo se vale para as demais classificações e abordagens para uso do algoritmo PageRank em folksonomias, como discutido no capítulo 2.

c) Distribuição:

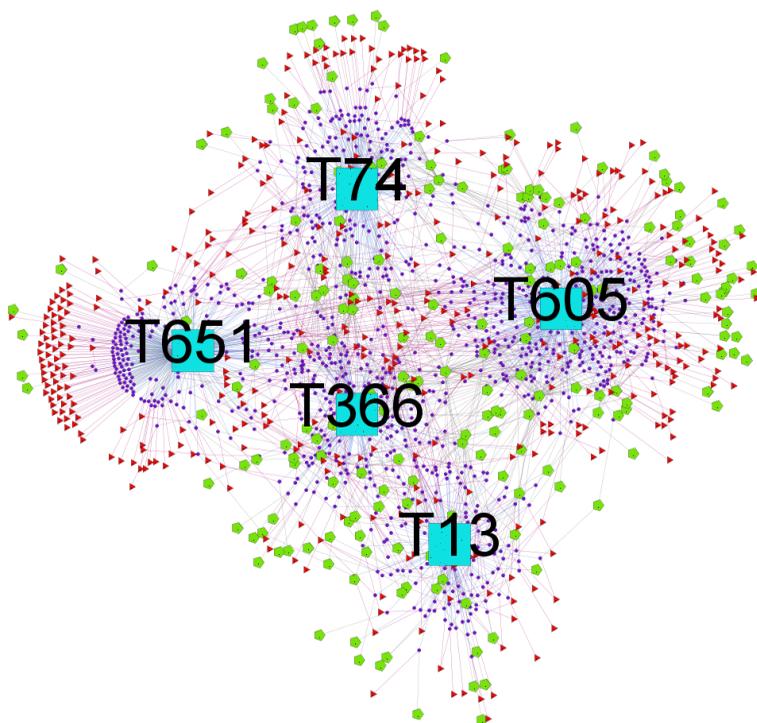
Dando continuidade à análise do modelo, os agrupamentos e comunidades podem ser identificados no grafo de acordo com as técnicas de distribuição. Dentre as alternativas para a distribuição Mika (2007) indica a centralidade local (*betweenness centrality*), ao qual o cálculo permite distinguir os termos em geral ou específico. Nesse caso, um termo geral está conectado a um número maior de nós de diversos

grupos, enquanto termos específicos, normalmente, estão conectados mais fortemente a nós mais próximos.

A filtragem das *tags* foi realizada pelo software Gephi com uso dos filtros: Partição, para separar apenas as *tags*; Intervalo, para filtrar apenas as *tags* com maior grau de centralidade local, até alcançar apenas as 5 de maior grau; em seguida foi adicionado o filtro para buscar a rede de vizinhos em nível 2. O resultado do subgrafo é apresentado na Figura 28, destacando as *tags* em quadrados com a cor azul.

Para identificar os agrupamentos que ocorrem a partir das *tags*, a Tabela 4 destaca os cinco termos de maior centralidade local em ordem. Os respectivos nós são comparados a classificação da *tag* se considerar o grau e o PageRank do nó na rede, apresenta também a quantidade de usuários e recursos alcançados.

Figura 28 – *Tags* selecionadas por grau de centralidade local e seus respectivos nós vizinhos.



Fonte: do autor.

Tabela 4 – Dados das cinco *tags* com maior centralidade local.

<i>Tag</i>	Degree	PageRank	Qnt User	Qnt Recursos
T605	1°	1°	109	183
T74	4°	4°	54	107
T13	256°	12°	51	44
T366	5°	5°	68	64
T651	3°	3°	29	164

Fonte: do autor.

A fim de identificar a quantidade de usuários e recursos que utilizam a *tag*, foi realizado um filtro da Rede Ego de cada nó, e, a partir desse, os vizinhos em nível dois de profundidade. Assim, é apresentada a rede de vizinhos de cada *tag*. Para a contagem, bastou adicionar um novo filtro de partição e, então, no menu “contexto” do software Gephi, são apresentados os números de nós com tal tipo.

Podemos observar que nós com maior centralidade local são mais propensos a se relacionar com diferentes grupos, sendo identificados por Mika (2007) como termos gerais. Na abordagem das folksonomias para avaliação de ideias, tais *tags* representam uma maior importância para casos de pesquisa, pois os resultados afetariam um maior número de ideias e também satisfaria a vontade de um maior grupo de usuários. Porém, para uma pesquisa mais focada a uma determinada área ou conceito, os termos mais específicos, apresentam critérios avaliados mais reservadamente a um grupo.

A seguir, para as demais análises, são destacados os recursos com maior PageRank dentro do grafo completo. Em um primeiro momento, as estatísticas não foram recalculadas, assim se manteve os números do grafo utilizado nas análises anteriores.

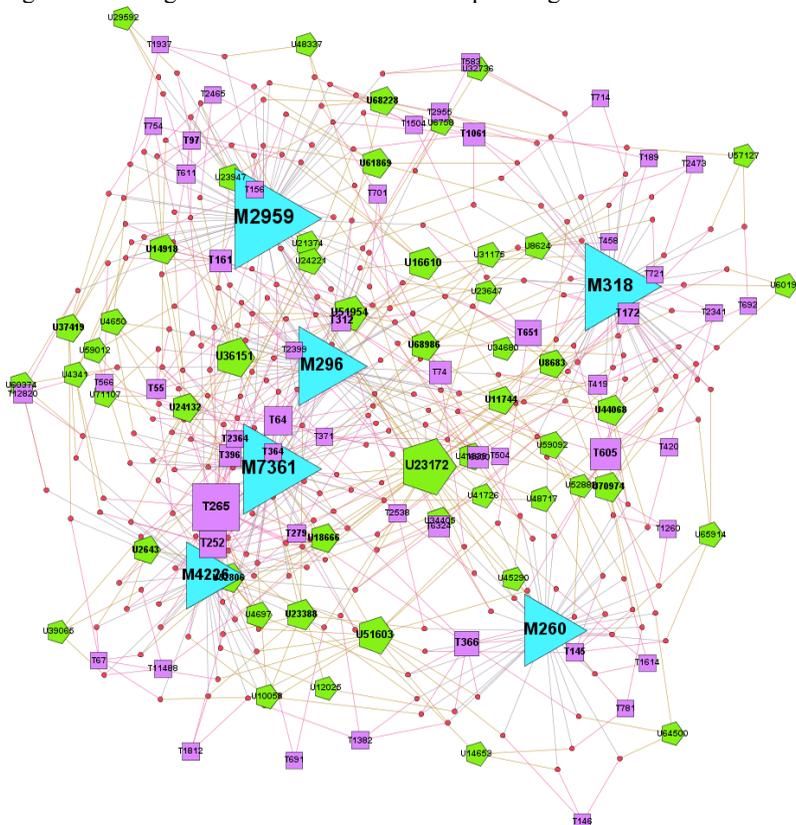
Nesse novo subgrafo, a partir do software Gephi, foram aplicados os filtros: “Partição”; para selecionar apenas os nós da classe “movie”; “Intervalo de partição”, com a escala do PageRank, fechando até alcançar apenas os filmes com maior grau, devido à proximidade, foram mantidos os 6 primeiros; em seguida foram selecionados os respectivos vizinhos com auxílio do filtro “Rede de Vizinhos” em nível 2, assim alcançando as atribuições, *tags* e usuários conectados.

A fim de reduzir e fechar o grupo ainda mais, foi realizada a filtragem com o filtro K-core, onde K=3. Então, foram alcançados cerca de 400 nós a serem utilizados para as análises a seguir.

O subgrafo que representa a comunidade dos seis recursos mais bem classificados, segundo o algoritmo PageRank, possui: 296 atribuições; 53 *tags*; 52 usuários. O subgrafo teve as estatísticas recalculadas pelo software Gephi para análise da comunidade gerada. A representação com o algoritmo de Hu (2005) é demonstrada na Figura 29, destacando os recursos em triângulos com a cor azul.

Na Figura 29, o grafo apresenta o tamanho dos nós de recursos e usuários regulados pelo algoritmo PageRank. O tamanho dos nós das *tags* consideram a centralidade local e as atribuições foram reduzidas para simplificar a visualização. A partir desse grafo, a Tabela 5 destaca algumas *tags* atribuídas.

Figura 29 – Subgrafo dos recursos destacados pelo PageRank entre 2006-2008.



Fonte: do autor.

Tabela 5 – Relação das *Tags* por centralidade local.

<i>Tag</i>	No conjunto	No grupo	Usuários	Recursos
T265*	24°	1°	16	3
T605	1°	2°	10	3
T64*	11°	3°	11	2
T252*	22°	4°	11	2
T651	4°	5°	3	5
T74	2°	9°	4	3
T366	3°	6°	9	2
T1061	5°	8°	5	2

Fonte: do autor.

Na Tabela 5, além das *tags* que apresentam, em ordem, a classificação da centralidade em relação ao conjunto de dados completo – T605, T74, T366, T651 e T1061 –, foram incluídas as *tags* com maior centralidade local, considerando o subgrafo formado pelos seis filmes – T265, T64 e T252 –, somando oito *tags* a serem observadas.

Pois, a boa classificação no conjunto completo, não representa a importância da *tag* para a subcomunidade formada pelos recursos selecionados. Assim, ao considerar além das etiquetas mais bem classificadas pela centralidade do conjunto, selecionar também as etiquetas com maior centralidade local, é possível abranger melhor os diferentes nós da subcomunidade, reduzindo as distâncias entre os recursos e usuários associados.

Assim, é observado que a classificação no grupo é diferente para o conjunto, o que representa uma maior abrangência em relação ao número de usuários e recursos próximos da *tag*. Sendo que, a *tag* com maior destaque – T245 – ocupa a 24° colocação no conjunto. Porém, dentro do grupo ela abrange 16 usuários e 3 dos 6 filmes analisados. Se mostrando um forte critério para esse grupo de usuários e recursos.

Uma análise semelhante foi realizada com os usuários, considerando o algoritmo PageRank em relação ao conjunto e em relação a subcomunidade, porém, sem diferenças significativas, apenas na ordenação. Se considerar a “*folkauthority*” por Côgo e Pereira (2014), a importância local de tais usuários é semelhante ao conjunto.

d) **Recomendação:**

A principal função da recomendação é analisar similaridades entre pares, formando grupos com preferências semelhantes. A análise de similaridades e recomendações observou a subcomunidade formada pelos usuários, *tag* e atribuições dos seis recursos mais bem ranqueados.

Tabela 6 – Coocorrência das principais *Tags* mais importante com os Recursos.

Movie/Tag	M260	M296	M318	M2959	M4226	M7361
T265	-	8	-	-	19	2
T605	4	3	13	-	-	-
T64	-	-	-	3	-	16
T252	-	-	-	-	11	10
T651	1	2	4	2	-	2
T74	3	1	-	2	-	-
T366	14	-	-	-	-	4
T1061	-	-	6	3	-	-

Fonte: do autor.

Primeiramente os dados foram cruzados em planilha e contabilizado o número de coocorrências. Para isso, os nós de atribuição foram filtrados e, então, verificadas as combinações dos nós: “movie” com “tag” (MT); “movie” com “user” (MU); “tag” com “user” (TU).

Na primeira análise do subgrafo Movie/Tag, a Tabela 6 apresenta o número de atribuições que as *tags* receberam em cada recurso. Se considerar que cada usuário atribui apenas uma vez cada *tag* ao recurso, é possível observar certa concordância nas atribuições.

Entre as associações analisadas, a etiqueta T265 que foi atribuída 19 vezes para o recurso M4226 obteve maior concordância entre os usuários. Outro destaque é a etiqueta T252 que foi atribuída 11 vezes para o recurso M4226 e 10 vezes para o recurso M7361, se mostrando uma forte característica na avaliação dos usuários, conforme Tabela 6.

A Tabela 7, que representa o subgrafo Movie/User apresenta as atribuições por usuários em cada um dos seis recursos. Os usuários observados tendem a marcar um grande número de *tags* por recurso.

Apesar do grande número de atribuições, é utilizada apenas uma nota ao recurso, devido à limitação do dataset. Considerando o nó de atribuição, a Tabela 8 apresenta essas notas marcadas aos recursos por cada usuário.

Tabela 7 – Atribuições pelos principais usuários para os recursos.

Movie/User	M260	M296	M318	M2959	M4226	M7361
U23172	10	1	12	8	7	6
U36151	-	9	-	3	6	6
U51603	11	-	-	3	10	-
U50846	-	-	7	-	7	-
U51954	-	3	3	4	-	4

Fonte: do autor.

Tabela 8 – Médias atribuídas pelos principais usuários para os recursos.

Movie/User	M260	M296	M318	M2959	M4226	M7361
U23172	4,0	4,5	4,5	4,5	4,0	4,5
U36151	-	5,0	-	5,0	5,0	4,5
U51603	3,0	-	-	4,0	5,0	-
U50846	-	-	5,0	-	5,0	-
U51954	-	4,0	3,0	5,0	-	5,0

Fonte: do autor.

Com as médias atribuídas por recursos, é possível identificar similaridades entre os usuários. A partir dos dados da Tabela 8, é calculada a similaridade das notas atribuídas pelo coeficiente de Pearson, os resultados são conferidos na Tabela 9.

Nesse caso, os usuários U23172 e U51954 apresentam maior similaridade em relação às notas atribuídas e aos recursos da subcomunidade, enquanto o usuário U51603 apresenta discordância com o usuário U23172.

Ao observar novamente o grafo completo, as *tags* analisadas são recorrentes pelos usuários, sendo aplicadas em diversos recursos fora da subcomunidade, conforme Tabela 10.

Tabela 9 – Similaridades das notas atribuídas pelos Usuários pelo coeficiente de Pearson.

User/User	U23172	U36151	U51603	U50846	U51954
U23172	1,0				
U36151	0,2	1,0			
U51603	-0,7	0,2	1,0		
U50846	-0,3	-0,2	0,2	1,0	
U51954	0,9	0,4	-0,5	-0,4	1,0

Fonte: do autor.

Tabela 10 - Uso das *Tags* pelos usuários no conjunto geral.

Tag/User	T265	T605	T64	T252	T651	T74	T366	T1061
U23172	3	22	8	2	5	4	-	37
U36151	1	6	-	-	-	-	4	-
U51603	2	3	2	1	3	11	4	-
U50846	1	-	2	2	-	-	-	-
U51954	-	9	3	-	-	24	-	-

Fonte: do autor.

Tabela 11 – Relação da similaridade entre as personomias dos cinco usuários mais destacados.

<i>User/User</i>	U23172	U36151	U51603	U50846	U51954
U23172	1,0				
U36151	0,1	1,0			
U51603	-0,4	0,0	1,0		
U50846	-0,3	-0,4	-0,4	1,0	
U51954	-0,1	0,0	0,9	-0,3	1,0

Fonte: do autor.

As repetidas atribuições de etiquetas por cada usuário gera uma “personomia”, como indicado por Cattuto *et al.* (2007). Por vezes, essas “personomias” combinam com as ocorrências de outros usuários, o que é possível calcular como um grau de similaridade. Tal similaridade pode servir, então, como informações para um sistema de recomendação entre os nós.

As atribuições das *tags*, usuários e os respectivos filmes foram filtrados pela “Rede Ego” e, então, calculadas a concordância entre tais atribuições. Para calcular essa semelhança foi escolhido o algoritmo de Pearson devido às ocorrências apresentar um peso diferenciado, ou seja, a quantidade de atribuições diferencia o grau de similaridade entre as atribuições. A Tabela 11 apresenta o resultado da semelhança entre as atribuições por usuário.

Nesse caso, os usuários U51603 e U51954 apresentam semelhanças positivas entre suas personomias. O usuário U50846 mantém semelhanças negativas, ou contrárias à comunidade.

Como observado por Xu *et al.* (2008), a semelhança em sistemas de folksonomias necessita considerar ainda uma técnica para poda. Considerando então apenas o grupo dos seis filmes com maior PageRank e os recursos associados, a Tabela 12 apresenta o número de concordâncias entre os usuários para o grupo.

Tabela 12 – Coocorrências de atribuições semelhantes para os 6 filmes do conjunto.

Movie/Tag	T265	T605	T64	T252	T651	T74	T366	T1061
U23172	1	1	2	2	-	-	-	-
U36151	1	1	-	-	-	-	1	-
U51603	1	1	-	1	1	1	1	-
U50846	1	-	-	1	-	-	-	-
U51954	-	-	2	-	-	2	-	-

Fonte: do autor.

Tabela 13 – Similaridade entre as atribuições dos usuários em relação ao conjunto de 6 filmes.

User/User	U23172	U36151	U51603	U50846	U51954
U23172	1,0				
U36151	-0,1	1,0			
U51603	-0,2	0,4	1,0		
U50846	0,5	0,1	0,3	1,0	
U51954	0,2	-0,4	-0,3	-0,3	1,0

Fonte: do autor.

Na primeira análise das personomias, apresentada na Tabela 11, havia pouca concordância entre os usuários selecionados. Mas, em relação à personomia dos 5 principais usuários para os 6 filmes a concordância entre os usuários é melhor detalhada, conforme Tabela 12.

Considerando os dados da Tabela 13, é possível calcular a similaridade entre os termos e entre os usuários. A Tabela 13 apresenta os resultados da correlação de Pearson.

Se no conjunto geral o usuário U50846 não apresentava concordância com os demais usuários, dentro do grupo há uma predominância por concordar positivamente com as atribuições. Já o usuário U51954 concorda apenas com algumas marcações do U23172. Outras relações de similaridade podem ser realizadas do mesmo modo, como a similaridade entre *tags* e recursos. Nesse caso, a similaridade aponta a concordância de *tags* atribuídas para o conjunto de filmes.

e) Análise temporal:

Por último é necessário relacionar a propriedade “*timestamp*” aos algoritmos e técnicas apresentados. A análise temporal permite identificar as variações e tendência do grafo em relação ao tempo.

Para essa observação, ainda considerando a comunidade formada pelos seis recursos com maior PageRank do conjunto geral e as 53 etiquetas atribuídas, é apresentada a análise em relação ao período de atribuições.

Inicialmente, a fim de verificar a importância dos nós por período, dentro da comunidade local, o PageRank foi recalculado por faixas de tempo, resultando em uma observação mais detalhada das tendências e considerando o LocalRanking, discutido por Gedikli (2013).

Para obter essa classificação, foi ativado no software Gephi a opção de coluna dinâmica com a propriedade “*timestamp*” e filtrados os nós com a classe “*tag*”. O resultado das 10 *tags* com melhor posição do

LocalRanking são apresentadas na Tabela 14, que relaciona também a posição dessas no decorrer dos três anos.

O destaque são as etiquetas T265, T64 e T97 que alcançaram um crescimento significativo em relação à posição. Compreender essas alterações da comunidade é essencial para os estudos da folksonomias. A partir da interação social, determinadas *tags*, recursos ou usuários perdem o engajamento e a importância na rede.

Ao considerar a propriedade tempo, é possível observar variações nessas constatações, sendo que, em determinados períodos algumas *tags* são mais utilizadas, recursos melhores pontuados, ou, então, *tags* que não tinham a menor representatividade, passam a se destacar. Assim, a mudança de comportamento aponta tendências e permite uma avaliação flexível ao cenário, conforme Landia *et al.* (2013).

A oscilação faz parte das folksonomias, por conta de que, criar um consenso entre os participantes está suscetível a mudanças do contexto e se mantém instável. Assim como no ingresso de usuários, pois agregam novas atribuições. Tal mudança é bastante benéfica ao contexto de avaliação de ideias, pois permite um resgate de ideias que foram rejeitadas em determinado momento, porém ganham valor por um determinado fato novo.

Portanto, a mudança de cenários, que poderia representar um problema para a avaliação de ideias, passa a ser uma linha de análise ao considerar as folksonomias e mudanças de contexto, por permitir que ideias voltem a se destacar.

Tabela 14 – Relação das 10 *tags* com maior destaque e a classificação por ano.

<i>Tags</i>	2006	2007	2008	Geral
T265	16°	9°	1°	1°
T252	15°	30°	28°	2°
T64	19°	15°	8°	3°
T172	17°	27°	5°	4°
T312	22°	17°	12°	5°
T605	20°	49°	9°	6°
T161	18°	38°	21°	7°
T366	21°	48°	7°	8°
T97	35°	29°	3°	9°
T1061	1°	13°	2°	10°

Fonte: do autor.

Tabela 15 - Atribuições de *tags* por filmes da interseção.

<i>Tag/movie</i>	T366	T605	T651
M1210	13	1	1
M1196	12	2	2
M260	14	4	1
M1270	4	2	1

Fonte: do autor.

Como conclusão das análises, ao considerar alguns critérios, é possível selecionar recursos não apenas pela popularidade ou pontuação, mas definir quais critérios são essenciais e, a partir desses obter um conjunto de recursos e suas notas. Para essa análise foram escolhidas três *tags* que apresentaram similaridade no cálculo do coeficiente de Pearson: T605, T366 e T651.

Após a escolha das *tags* foram filtrados os nós no software Gephi: Topologia “Rede Ego” em nível 2 para cada *tag*; Operador “Intercessão”; e, por fim, o filtro Partição para a classe “movie”. O resultado dos recursos da intercessão e número de atribuições da *tag* para o filme são apresentados na Tabela 15.

Em cada atribuição uma nota é atribuída ao recurso, portanto, para a análise, foi necessário calcular a média das atribuições para cada recurso em relação as *tags* escolhidas. Para cada recurso, a Tabela 16 apresenta o número de atribuições da etiqueta T366 por nota na escala de 0,5 a 5,0 e a média calculada. Semelhantemente foi realizado o cálculo para as demais etiquetas, o resultado é apresentado na Tabela 17.

A Tabela 17 apresenta: as médias obtidas em relação as atribuições de *tags* para cada recurso; a média considerando pesos semelhantes para as *tags* escolhidas; e a relação com a média geral de atribuições para os recursos.

Tabela 16 – Número de atribuições na escala de notas para o nó T366.

M/T366	0,5	1,0	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0	Média
M1210	-	-	-	-	1	2	2	2	2	4	4,0
M1196	-	-	-	-	-	-	2	4	1	5	4,4
M260	-	-	-	1	-	3	2	1	2	5	4,0
M1270	-	-	-	-	-	-	-	-	-	4	5,0

Fonte: do autor.

Tabela 17 – Média das *Tags*

<i>Tag/movie</i>	T366	T605	T651	Média	Média Geral
M1210	4,0	3,5	3,5	3,7	3,5
M1196	4,4	4,8	5,0	4,7	3,1
M260	4,0	3,9	3,0	3,6	3,3
M1270	5,0	4,5	4,0	4,5	3,0

Fonte: do autor.

Como observado na Tabela 17, os recursos M1196 e M1270, que possuíam média geral inferior, passaram a ter melhor pontuação na avaliação que considerou apenas algumas etiquetas.

4.4.3 Os Resultados

A análise do modelo considerou as técnicas identificadas pela mineração de folksonomias (Hotho, 2010) como: (a) Decomposição de grafos; (b) Classificação; (c) Distribuição; e (d) Recomendação; sendo ainda combinadas a fim de atingir os objetivos.

Uma técnica fundamental para explorar o conhecimento representado na rede foi a decomposição do hipergrafo. Ao qual permite a continuidade da análise para as demais técnicas, devido os diferentes tipos de nós e o tamanho do hipergrafo representar uma rede complexa de ser observada.

Em relação a isso, por apresentar uma clara definição da “atribuição” como um nó intermediário, responsável por realizar a ligação entre usuários, recursos e *tags*, o modelo persiste as informações de data e nota como propriedades nesse nó. Qualquer decomposição que resulte na eliminação do nó consiste na descoberta de novos conhecimentos em detrimento de uma redução das informações. O grafo resultante não permite o regresso ao hipergrafo, pois as informações reduzidas são representadas por pesos nas conexões, geralmente.

Ao considerar a classificação dos nós no grafo, uma composição de estatísticas, algoritmos e decomposições possibilitam verificar a popularidade (*degree*), importância (*PageRank*) e alcance (*betweenness centrality*). Essa observação foi fundamental, pois permitiu evidenciar, por exemplo, que o grau do nó não representa sua importância ou suas características dentro de um grupo local. A partir disso, se estabelece a primeira constatação:

1ª DIRETRIZ: Etiquetas populares não são, necessariamente, um consenso de critérios pela comunidade.

Como observado na Figura 27, ao qual uma *tag* popular não representa o consenso para os usuários. Essa constatação corrobora com a citação de Gruber (2008): “*popularidade é uma medida de qualidade e não uma medida de autenticidade*”.

A Web semântica social busca justamente essa diferenciação das métricas de participação social ao abordar a simbiose do humano e máquina. Nesse conceito, o conhecimento necessita ser contextualizado, o que é possível com a representação do conhecimento.

Portanto, para um Sistema de Gestão de Ideias identificar o conhecimento coletivo as alternativas são: utilizar os algoritmos de PageRank discutidos na literatura por Hotho (2010), a fim de identificar a importância do nó na rede; ou, então, o alcance pela característica de comprimento do caminho discutidos por Cattuto *et al.* (2007). Em relação as etiquetas a centralidade local se mostrou mais adequada por abranger um maior número de usuários e recursos.

Outro aspecto a ser considerado é a fragmentação das etiquetas. Apesar de evidenciar uma complexidade das folksonomias, a fragmentação expressa um conhecimento enriquecedor de significados para a base de conhecimento. Portanto é definida a constatação:

2ª DIRETRIZ: A fragmentação das etiquetas distingue a aplicação dos critérios em específicos ou locais.

O mais adequado para evidenciar os critérios de uma comunidade é a observação da rede de conhecimento envolvida nas avaliações. Isso considera o consenso, a semântica e a qualificação atribuída.

Ou seja, todas as etiquetas são recomendações de características a serem consideradas, mas o que é imprescindível para a observação é esclarecer os motivos das atribuições pela comunidade, considerando se esses critérios realmente importam para o contexto.

Para esse fim, os agrupamentos e distribuições dos conjuntos são percebidos como mais adequados para análise, pois permitem perceber a similaridade entre as atribuições e indicar um direcionamento geral ou específico das etiquetas.

Assim, identificamos pelo estudo que, em cada subcomunidade, as etiquetas podem carregar um valor positivo, ou um valor negativo. Essa percepção é possível a partir de uma estruturação por agrupamento. Exemplificando: Ao identificar as etiquetas atribuídas aos seis principais filmes do conjunto, percebemos a pontuação positiva para os critérios de: (a) Roteiro, se considerar as *tags* como “based on book” (T1061),

“surreal” (T64), “memory” (T252), “time travel” (T13); (b) Temática, como “sci-fi” (T366), “action” (T74), “comedy” (T55); e (c) Elenco/Pessoas, tais como “Edward Norton” (T97), “Johnny Depp” (T59), “Woody Allen” (T61), “Quentin Tarantino” (T312). Ao detectar avaliações negativas, o agrupamento permite identificar o motivo das ideias/recursos serem rejeitados pela subcomunidade. Nesse caso, as *tags* atribuídas negativamente estão associadas a cenas como: “nudity” (T93), “violence” (T161) e “murder” (T1023).

Portanto, para a subcomunidade definiríamos os critérios em: roteiro, temática, elenco/pessoas e cenas, ao qual, as respectivas etiquetas seriam agrupadas semanticamente no sistema de gestão de ideias por vocabulários formais para a inferência do conhecimento.

Já na distribuição, os agrupamentos e a observação das subcomunidades apresentam pontos importantes a serem considerados, por exemplo, se um conhecimento é genérico para a comunidade, ou específico para uma subcomunidade, semelhante a discussão por Mika (2007). Do mesmo modo, as personomias possibilitam identificar quais etiquetas que determinados usuários estão atribuindo e o modo que utilizam: positivamente ou negativamente. Essas etiquetas em concordância com os demais usuários demonstram o conhecimento coletivo empregado.

Essa breve análise já destaca a importância que uma análise local pode representar. Portanto, é possível definir que:

3ª DIRETRIZ: A importância local identifica mais adequadamente o conhecimento da subcomunidade.

O conhecimento coletivo não é uma medida de popularidade, conforme a primeira constatação. Assim, ao observar especificamente o comportamento de uma subcomunidade, é possível detectar melhor os padrões de interação social, similar ao discutido por Cattuto *et al.* (2007) como o “pequeno mundo”.

Por uma subcomunidade é compreendida a composição dos usuários, recursos, etiquetas e as atribuições dentro de um subgrafo. Essa estrutura distingue, mais adequadamente, o comportamento de determinados usuários, em relação a recursos específicos e o padrão de atribuição, como identificado por Ghosh *et al.* (2011).

Assim, enquanto para uma subcomunidade um recurso é considerado negativo por possuir certa característica, para o mesmo grupo de usuários a característica pode ser considerada positiva em outros recursos. Ou então, identificar a contradição entre grupos, pois a

qualificação expressa a preferência dos indivíduos. Essa preferência também caracteriza uma recomendação mais adequada. Ou seja:

4ª DIRETRIZ: A recomendação contextualizada pela semântica e notas resultam em similaridades mais precisas.

Como observado na segunda constatação, uma *tag* pode ser atribuída por usuários de forma positiva, como também negativa. Isso permite contextualizar uma avaliação a partir não só da nota atribuída ao recurso, mas também dos critérios apontados.

No que se refere às notas atribuídas, com base nas coocorrências, habilitam a descoberta de conhecimentos para definir recomendações. A análise dos coeficientes de similaridade, calculados dentro das subcomunidades, garantem um conhecimento mais confiável e propício para os agentes humanos por considerar a semântica.

Nesse caso, a observação de um usuário sobre determinado recurso é identificada por sua personomia e por sua respectiva avaliação nos recursos. Assim, favorece o desenvolvimento de sistemas de conhecimento coletivo para recomendação, pois indica uma contextualização das preferências.

A recomendação se estende para os subgrafos o que permite ao sistema sugerir: (a) ideias para um determinado contexto, expresso por um conjunto de *tags*; (b) etiquetas ainda não atribuídas para determinadas ideias; (c) usuários como avaliadores para certos recursos; (d) etiquetas para complementar as personomias dos usuários; (e) etiquetas com significados semelhantes; (f) recursos para os usuários avaliar; (g) ideias semelhantes que poderão ser enriquecidas mutuamente, talvez baseadas em suas deficiências/qualidades; entre outras possibilidades.

Ao considerar a recomendação de ideias com base em determinados contextos, é possível então acompanhar as mudanças estratégicas e recuperar ideias rejeitadas. Logo:

5ª DIRETRIZ: A adaptação ao contexto organizacional é baseada na variação temporal, tanto na descoberta como na representação do conhecimento.

Todas as observações anteriores estão propensas à análise temporal. Considerando que uma associação é dependente da existência do nó de atribuição, ao habilitar a dinâmica temporal, certas associações só passam a existir em determinada faixa de tempo. Esse processo permite analisar a evolução do grafo tal como a interação social.

O conhecimento coletivo é dependente dessa interação social e, assim, a adaptação ao contexto organizacional é baseada nessa evolução, tanto na descoberta como na representação do conhecimento.

Consequentemente, a variação das atribuições corrobora com o objetivo de resgatar recursos rejeitados, pois permite a flexibilização das avaliações, revelar tendências e direcionar pesquisas de acordo com as atribuições sociais. O conceito de flexibilização da avaliação seguindo o contexto, é percebido como uma necessidade pelos estudos de Brix (2015) e Sandström e Björk (2010) para a Inovação.

4.5 CONSIDERAÇÕES DO CAPÍTULO

Os artefatos e constatações discutidos, permitem evidenciar os objetivos da pesquisa. O primeiro artefato resultante da pesquisa, segundo a *design science*, se caracteriza como um método, pois “*apresenta um conjunto de passos usados para executar uma tarefa*” (Dresch *et al.*, 2015).

Como fundamento do método de avaliação por folksonomias, são abordados os processos da Engenharia do Conhecimento como: aquisição; representação e armazenamento; raciocínio e extração do conhecimento. A partir desse conhecimento é necessária a definição pela organização baseada no contexto.

Por considerar uma forma de detecção de critérios a partir do conhecimento coletivo, o método proposto se diferencia dos métodos comuns que consideram índices de popularidade ou modelos de formulários com critérios pré-estabelecidos. Em sistema de gestão de ideias, isso representa o aproveitamento semântico e qualitativo do conhecimento coletivo.

Para esse fim, o método consiste em perceber o conhecimento aplicado a partir de etiquetas atribuídas livremente. Assim, particularmente no *front end* da inovação, quando as informações são ainda vagas e incertas, o método permite que a comunidade expresse sua avaliação por etiquetas, notas, e, então, obtém uma definição de critérios a serem compreendidos.

Além do método, foi necessário o desenvolvimento de um segundo artefato, um modelo que permitisse a sua “*representação do espaço de solução*” (Dresch *et al.*, 2015). O modelo considera a representação do conhecimento coletivo em uma estrutura típica das folksonomias, um hipergrafo (Lohmann e Díaz, 2012).

Pela orientação para armazenamento em grafo, o modelo foi concebido como um hipergrafo não uniforme, ao qual as informações são inseridas em nós, para evitar o uso das arestas. Esse modelo se mostrou adequado para a mineração de folksonomias, permitindo inferir o conhecimento empregado pela comunidade.

Outras propostas de análise a serem consideradas em um Sistema de Gestão de Ideias, a partir desse método, são: (a) Apontar as deficiências ou carência de conhecimentos para determinada área; (b) Reduzir os erros a partir de percepção de avaliações por critérios subjetivos e pelo histórico do processo; (c) Definir pesos para usuários por sua expertise, a *folkauthority* apresentada por Côgo e Pereira (2014).

No próximo capítulo são discutidas as considerações finais apontando as demais pesquisas futuras, limitações e principais resultados em relação aos objetivos.

5 CONSIDERAÇÕES FINAIS

Este capítulo apresenta as contribuições da dissertação e as recomendações para trabalhos que possam ser desenvolvidos posteriormente.

5.1 REVISÃO DO PROBLEMA E DA PROPOSTA

Os desafios da Web semântica social se apresentam na simbiose do conhecimento gerado por humanos e computado por máquina para inferir novas descobertas. A aquisição desse conhecimento para o domínio de avaliação de ideias, em especial no *front end* da Inovação, representa a contribuição científica dessa dissertação.

Nos Sistemas de Gestão de Ideias, a principal forma de avaliação social são dados de popularidade. Porém, tais dados não expressam semanticamente o conhecimento empregado pelos usuários, como ressaltado por Gruber (2008).

Enquanto a popularidade não evidencia uma medida de qualidade, formulários fechados podem não ser capazes de abranger critérios adequados para o contexto, mesmo porque as características das ideias no FEI não permitem uma validação confiável, mesmo por especialistas. Mas, baseado na leitura dos artigos sobre o conhecimento coletivo, uma forma de responder a essa necessidade é perceber o conhecimento coletivo nas avaliações. A alternativa encontrada foi o uso de atribuições sociais, as folksonomias.

A partir das Folksonomias é possível identificar etiquetas relacionadas aos recursos. As etiquetas expressas juntamente com uma avaliação qualitativa, permitem identificar os motivos examinados pelo usuário ao pontuar um recurso.

O princípio do método considera uma base de ideias ao qual os usuários expressam semântica e qualitativamente o conhecimento empregado em uma avaliação. Esse conhecimento é então representado por um modelo para o armazenamento; a agregação das interações gera um conhecimento coletivo a ser analisado por técnicas de mineração de folksonomias; e o produto é, então, um conjunto de características a serem consideradas pela organização como critérios.

Ou seja, em um primeiro momento, se as ideias não possuem critérios claros a serem considerados para a avaliação, o modelo permite perceber esses critérios a partir da interação social e semântica. Além disso, a interação temporal tende a emergir folksonomias cada vez mais

consistentes. Assim, a seleção de ideias, parte do princípio de critérios mais claros e destacados pela comunidade, convergindo para o cenário, estratégia e cultura organizacional. Logo, o processo de seleção poderá se valer desses critérios destacando quais os requisitos necessários para o contexto atual ou uma perspectiva futura. Desse modo, a análise permitiu garantir que o método é adequado às características apresentadas pelo FEI.

Assim, considera-se alcançado o objetivo de “*propor um método de identificação e análise de critérios para avaliação de ideias a partir do conhecimento coletivo em um sistema de gestão de ideias*”.

Em relação ao propósito de atender a cada objetivo específico, conclui-se que:

- a) **Identificar critérios de análise de ideias:** a aquisição do conhecimento coletivo, baseado na interação social, permitiu identificar critérios de análise.
- b) **Analisar o conhecimento coletivo na avaliação de ideias:** a partir da aquisição e representação, foi possível analisar o conhecimento com técnicas de mineração de folksonomias.
- c) **Avaliar ideias em diferentes contextos:** os resultados propiciam uma avaliação mais adequada por atribuição social para diferentes contextos, pois, o método habilita um nível semântico e social conveniente para o resgate de ideias.

5.2 GENERALIZAÇÃO

Já por usar um conjunto com informações de filmes e não de ideias, o método propicia a sua generalização para outros recursos. Apesar do direcionamento da pesquisa considerar os problemas encontrados na avaliação de ideias no *front end* da Inovação, como generalização do artefato, outros sistemas que consideram a avaliação social podem perceber benefícios semelhantes com o método.

Por sistemas de avaliação social podem ser abrangidos os recursos como filmes, vídeos, músicas, notícias, textos, documentos, pesquisas de satisfação, entre outros, respeitando as especificidades de cada recurso.

Em todos esses casos a semântica das avaliações não é explícita, e não é possível identificar o motivo de cada usuário avaliar positiva ou

negativamente um recurso. Assim, tais sistemas poderiam valer do método e do modelo a fim de atingir resultados semelhantes.

5.3 CONTRIBUIÇÕES

Durante o mestrado foram publicados artigos científicos que auxiliaram na comunicação científica e rede de contatos acadêmicos. A partir das pesquisas realizadas para essa dissertação, o Quadro 4 apresenta as publicações relacionadas.

Quadro 4 – Publicações relacionadas a dissertação.

PUBLICAÇÕES	PALAVRAS-CHAVE
ROCHADEL, W.; DANDOLINI, G. A.; SOUZA, J. A.. Vínculos Externos no <i>Front End</i> da Inovação. In: 8º Congresso do Instituto Franco-Brasileiro de Administração de Empresas, 2015, Gramado/RS.	Inovação aberta; <i>Front End</i> ; Revisão; Ideias.
ROCHADEL, W.; SILVA, T. R.; GONÇALVES, A. L.; DANDOLINI, G. A.; SOUZA, J. A.. Modelo para Representação do conhecimento em um sistema de recomendação de avaliadores. In: 12th CONTECSI International Conference on Information Systems and Technology Management, 2015, Sao Paulo/SP.	Sistemas de Recomendação; revisores ad hoc; Engenharia do Conhecimento; mineração de dados; Lattes.
ROCHADEL, W.; VALDATI, A. B.; DANDOLINI, G. A.; SOUZA, J. A.. Inteligência para a Inovação na Avaliação e Seleção de ideias. In: IV International Conference on Design, Engineering, Management for innovation, 2015, Florianópolis/SC.	Inovação; Gestão de ideias; Avaliação de Ideias; Seleção de Ideias; Revisão Sistemática.
ROCHADEL, W.; DANDOLINI, G. A.; SOUZA, J. A.. Ontologias para a Gestão de Ideias na Inovação Aberta. In: IV International Conference on Design, Engineering, Management for innovation, 2015, Florianópolis/SC.	Inovação Aberta; Ontologia; Web Semântica; Gestão de Ideias; Engenharia do Conhecimento.
ROCHADEL, W.; DANDOLINI, G. A.; SOUZA, J. A.. Knowledge representation in collective ideation. <i>Strategic Design Research Journal</i> , Aceito para publicação em Outubro de 2015.	Open Innovation; Open Design; Idea Management; Knowledge Engineering.

Fonte: do autor.

5.4 LIMITAÇÕES

Por se tratar de uma proposta de método, a primeira limitação foi encontrar um conjunto de dados próprios para a análise.

O autor Adam Westerski disponibilizou o conjunto de dados extraídos da comunidade de geração de ideias “UbuntuBrainstorming”. O conjunto foi utilizado nos artigos Poveda *et al.* (2012) e Westerski *et al.* (2013). Porém, não apresenta folksonomias, então não pode ser aproveitado para as análises.

Assim, é sugerida como pesquisa futura a verificação com dados reais em diferentes domínios, mesmo por grupos fechados.

5.5 RECOMENDAÇÕES E TRABALHOS FUTUROS

Durante as fases criativas do FEI, ideias são descritas brevemente, sem informações adicionais e apresentam um alto nível de incerteza.

Como continuidade da pesquisa na atividade de seleção são sugeridas outras recomendações e análises a partir do método, como por exemplo: caso a *tag* escolhida não tenha uma nota atribuída ao recurso, inferir um valor ou calcular a incerteza da avaliação. Outra forma seria a recomendação de *tag* para os usuários avaliarem outros recursos do mesmo modo e, assim, criando um conjunto de critérios a partir do conhecimento coletivo para mais recursos.

Como principal perspectiva percebida para continuidade da pesquisa, é voltar a atenção para os aspectos de seleção, também abordado por outros autores, como cadeias de Markov, inferências, redes de incerteza. A complexidade dos hipergrafo pode colaborar nesses casos.

REFERÊNCIAS

- ALEXANDER, S. Researching networked learning generatively. **Proceedings of the 7th international Conference on Networked Learning**, 2008. ISSN 978-1-86220-206-1.
- ALVES, H.; SANTANCHÈ, A. Folksonomized ontology and the 3E steps technique to support ontology evolution. **Web Semantics: Science, Services and Agents on the World Wide Web**, v. 18, n. 1, p. 19-30, 1// 2013. ISSN 1570-8268. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S1570826812000789> >.
- ALVES, J. et al. Creativity and innovation through multidisciplinary and multisectoral cooperation. **Creativity and Innovation Management**, v. 16, n. 1, p. 27-34, 2007. ISSN 1467-8691.
- ANDRADE, R. P. D. A construção do conceito de incerteza: uma comparação das contribuições de Knight, Keynes, Shackle e Davidson. **Nova Economia**, v. 21, p. 171-195, 2011. ISSN 0103-6351. Disponível em: < http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-63512011000200001&nrm=iso >.
- AZNAR, G. **Idées-100 techniques de créativité pour les produire et les gérer**. Editions Eyrolles, 2011. ISBN 2212860366.
- BACHMAN, M. **GraphAware: Towards Online Analytical Processing in Graph Databases**. 2013. 149 (Master). Department of Computing, Imperial College London, London.
- BACHMAN, M. Modelling Data in Neo4j: Qualifying Relationships. 24 de outubro 2014. Disponível em: < <http://graphaware.com/neo4j/2013/10/24/neo4j-qualifying-relationships.html> >. Acesso em: 29 de novembro de 2015.
- BAEZ, M.; CONVERTINO, G. Innovation cockpit: A dashboard for facilitators in idea management. ACM 2012 Conference on Computer Supported Cooperative Work Companion, CSCW'12, 2012, Seattle, WA. p.47-48.
- BAILEY, B. P.; HORVITZ, E. What's your idea? A case study of a grassroots innovation pipeline within a large software company. 28th Annual CHI Conference on Human Factors in Computing Systems, CHI 2010, 2010, Atlanta, GA. p.2065-2074.

BALANCIERI, R. **Um Método baseado em ontologias para explicitação de conhecimento derivado da análise de redes sociais de um domínio de aplicação**. 2010. (Doutorado). Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina, Centro Tecnológico, Florianópolis, SC.

BALOH, P. et al. Roles of information technology in distributed and open innovation process. 13th Americas Conference on Information Systems, AMCIS 2007, 2007, Keystone, CO. p.1505-1514.

BAREGHEH, A.; ROWLEY, J.; SAMBROOK, S. Towards a multidisciplinary definition of innovation. **Management Decision**, v. 47, n. 8, p. 1323-1339, 2009. Disponível em: < <http://www.emeraldinsight.com/doi/abs/10.1108/00251740910984578> >.

BASTIAN, M. et al. **LinkedIn skills: large-scale topic extraction and inference**. Proceedings of the 8th ACM Conference on Recommender systems. Foster City, Silicon Valley, California, USA: ACM: 1-8 p. 2014.

BASTIAN, M.; HEYMANN, S.; JACOMY, M. Gephi: an open source software for exploring and manipulating networks. **ICWSM**, v. 8, p. 361-362, 2009.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic Web. **Scientific american**, v. 284, n. 5, p. 28-37, 2001. ISSN 0036-8733.

BESSANT, J.; TIDD, J. **Inovação e Empreendedorismo: Administração**. Bookman, 2009. ISBN 8577805115.

BETTONI, M.; BERNHARD, W.; BITTEL, N. Collaborative solutions quick&clean: The SFM method. 14th European Conference on Knowledge Management, ECKM 2013, 2013, Kaunas. p.44-51.

BETTONI, M. et al. Idea management by role based networked learning. **11th European Conference on Knowledge Management, ECKM 2010**, Famalicao, p. 107-116, 2010. ISSN 20488963 (ISSN); 9781906638719 (ISBN). Disponível em: < <http://www.scopus.com/inward/record.url?eid=2-s2.0-81555201807&partnerID=40&md5=c9381eeecb2d77c5e75c6c0d256c503d> >.

BJÖRK, J.; BOCCARDELLI, P.; MAGNUSSON, M. Ideation capabilities for continuous innovation. **Creativity and innovation management**, v. 19, n. 4, p. 385-396, 2010. ISSN 1467-8691.

BORGATTI, S. P.; EVERETT, M. G.; FREEMAN, L. C. Ucinet. In: (Ed.). **Encyclopedia of Social Network Analysis and Mining**: Springer, 2014. p.2261-2267. ISBN 1461461693.

BRAGA, M. C. G. Diretrizes para o design de mídias em realidade aumentada: Situar a aprendizagem colaborativa online. 2012.

BREM, A.; VOIGT, K.-I. Integration of market pull and technology push in the corporate front end and innovation management—Insights from the German software industry. **Technovation**, v. 29, n. 5, p. 351-367, 2009. ISSN 0166-4972. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0166497208000898> >.

BRIGNOLI, J. T. **Um modelo para suporte ao raciocínio diagnóstico diante da dinâmica do conhecimento sobre incertezas**. 2013. (Doutorado). Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina, Centro Tecnológico, Florianópolis, SC.

BRIX, J. The performance-improving benefits of a radical innovation initiative. **International Journal of Productivity and Performance Management**, v. 64, n. 3, p. 356-376, 2015. ISSN 17410401 (ISSN). Disponível em: < <http://www.scopus.com/inward/record.url?eid=2-s2.0-84923377373&partnerID=40&md5=06d2eb4412084a3fcd8ee74c45721728> >.

BRUN, E.; SAETRE, A. S.; GJELSVIK, M. Classification of ambiguity in new product development projects. **European Journal of Innovation Management**, v. 12, n. 1, p. 62-85, 2009. ISSN 1460-1060.

BUCHELE, G. T. **Adoção de Métodos, Técnicas e Ferramentas para Inovação: um levantamento em organizações catarinenses**. 2015. (Mestrado). Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina, Centro Tecnológico, Florianópolis, SC.

BUNGE, M. **Emergence and convergence: Qualitative novelty and the unity of knowledge**. University of Toronto Press, 2003. ISBN 0802088600.

BURKE, R. Hybrid recommender systems: Survey and experiments. **User Modeling and User-Adapted Interaction**, v. 12, n. 4, p. 331-370, Nov 2002. ISSN 0924-1868. Disponível em: < <Go to ISI>://WOS:000179499600001 >.

BURRELL, G.; MORGAN, G. **Sociological paradigms and organisational analysis**. london: Heinemann, 1979.

CANTADOR, I.; BRUSILOVSKY, P.; KUFLIK, T. Second workshop on information heterogeneity and fusion in recommender systems (HetRec2011). RecSys, 2011, Chigado, USA. p.387-388.

CANTADOR, I.; KONSTAS, I.; JOSE, J. M. Categorising social tags to improve folksonomy-based recommendations. **Web Semantics: Science, Services and Agents on the World Wide Web**, v. 9, n. 1, p. 1-15, 3// 2012. ISSN 1570-8268. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S1570826810000685> >.

CAPES, M. Documento de área 2009. **TRIÊNIO**, v. 2007, p. 09, 2009.

CARBONE, F.; CONTRERAS, J.; HERNÁNDEZ, J. Enterprise 2.0 and semantic technologies: A technological framework for open innovation support. **Proceedings of the European Conference on Knowledge Management, ECKM**, p. 191-199, 2010. Disponível em: < <http://www.scopus.com/inward/record.url?eid=2-s2.0-84872652304&partnerID=40&md5=d44a2b750f87dd135e24efed558a32c7> >.

CATTUTO, C. et al. **Time-varying social networks in a graph database: a Neo4j use case**. First International Workshop on Graph Data Management Experiences and Systems. New York, New York: ACM: 1-6 p. 2013.

CATTUTO, C. et al. Network properties of folksonomies. **AI Communications**, v. 20, n. 4, p. 245-262, 2007. ISSN 09217126 (ISSN). Disponível em: < <http://www.scopus.com/inward/record.url?eid=2-s2.0-36049006152&partnerID=40&md5=04ea7cfd26d35f213ae3ffba2ad4b7d> >.

CECI, F. Um Modelo semi-automático para a construção e manutenção de ontologias a partir de bases de documentos não estruturados. 2010.

CHAKRAVARTY, S. Cost-benefit analysis. **The New Palgrave: Economic Development**, Macmillan Press, New York, p. 76-84, 1989.

CHELVIER, R. et al. Decision making with a random walk in a discrete time markov chain. Proc. of the 6th Vienna Conference on Mathematical Modelling, ARGESIM Report, 2011. p.2178-2183.

CHEN, J. et al. **Research on product open innovation design based on knowledge**. Advanced Materials Research. 230-232: 819-823 p. 2011.

CHESBROUGH, H. W. **Open Innovation: The New Imperative for Creating and Profiting from Technology**. Harvard Business School Press, 2003. ISBN 9781578518371. Disponível em: < <http://books.google.com.br/books?id=4hTRWStFhVgC> >.

CHIBÁS, F. O.; PANTALEÓN, E. M.; ROCHA, T. A. GESTÃO DA INOVAÇÃO E DA CRIATIVIDADE NA ATUALIDADE. 2013, v. 3, p. 12, 2013-08-02 2013. ISSN 1807-1600. Disponível em: < <http://www2.ifrn.edu.br/ojs/index.php/HOLOS/article/view/1082/678> >.

CHRISTENSEN, C.; RAYNOR, M. **The innovator's solution: Creating and sustaining successful growth**. Harvard Business Review Press, 2013. ISBN 1422196577.

CÔGO, F. R.; PEREIRA, R. We Have Good Information for You: Cognitive Authority and Information. **The Evolution of the Internet in the Business Sector: Web 1.0 to Web 3.0**, p. 191, 2014. ISSN 1466672633.

COOPER, R. G. Perspective: The Stage-Gate® Idea-to-Launch Process—Update, What's New, and NexGen Systems*. **Journal of Product Innovation Management**, v. 25, n. 3, p. 213-232, 2008. ISSN 1540-5885. Disponível em: < <http://dx.doi.org/10.1111/j.1540-5885.2008.00296.x> >.

CORMICAN, K.; O'SULLIVAN, D. Enterprise knowledge management. Proceedings of the 17th Annual Conference of the Irish Manufacturing Committee, 2000.

DE LIDDO, A. et al. Collective intelligence as community discourse and action. ACM 2012 Conference on Computer Supported Cooperative Work Companion, CSCW'12, 2012, Seattle, WA. p.5-6.

DE NOOY, W.; MRVAR, A.; BATAGELJ, V. **Exploratory social network analysis with Pajek**. Cambridge University Press, 2011. ISBN 1139501089.

DELUGACH, H. S.; HINKE, T. H. Using conceptual graphs to represent database inference security analysis. **Journal of Computing and Information Technology**, v. 2, n. 4, p. 291-307, 1994.

DOROW, P. F. **O Processo de Geração de Ideias para Inovação: Estudo De Caso Em Uma Empresa Náutica**. 2013. (Mestrado). Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina, Centro Tecnológico, Florianópolis, SC.

DRESCH, A.; LACERDA, D. P.; JÚNIOR, J. A. V. A. **Design Science Research: Método de Pesquisa para Avanço da Ciência e Tecnologia**. Bookman Editora, 2015. ISBN 8582602995.

DUBINKO, M. et al. Visualizing tags over time. **ACM Transactions on the Web (TWEB)**, v. 1, n. 2, p. 7, 2007. ISSN 1559-1131.

EGC. Área de Concentração. 2004. Disponível em: < <http://www.egc.ufsc.br/index.php/egc/pos-graduacao/programa/areas-de-concentracao> >.

ELERUD-TRYDE, A.; HOOGE, S. Beyond the generation of ideas: Virtual idea campaigns to spur creativity and innovation. **Creativity and Innovation Management**, v. 23, n. 3, p. 290-302, 2014. ISSN 09631690 (ISSN). Disponível em: < <http://www.scopus.com/inward/record.url?eid=2-s2.0-84906058564&partnerID=40&md5=d0d43caaa25d771389e89f5c0338e7d2http://onlinelibrary.wiley.com/store/10.1111/caim.12066/asset/caim12066.pdf?v=1&t=i7penkrc&s=7c5f8aeb9a55bb90661161d87dbe962e034e0867> >.

ENGEL, K.; DIEDRICHS, E.; BRUNSWICKER, S. **Imp³rove: A European Project with Impact 50 Success Stories on Innovation Management**. European Union. Luxemburgo - Bélgica. 2010

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. From data mining to knowledge discovery in databases. **AI magazine**, v. 17, n. 3, p. 37, 1996. ISSN 0738-4602.

FERIOLI, M. et al. Understanding the rapid evaluation of innovative ideas in the early stages of design. **International Journal of Product Development**, v. 12, n. 1, p. 67-83, 2010. ISSN 1477-9056.

FERNANDES, R. F. **Uma proposta de modelo de aquisição de conhecimento para identificação de oportunidades de negócios nas redes sociais**. 2012. (Mestrado). Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina, Centro Tecnológico, Florianópolis, SC.

FERREIRA, J. M.; TACLA, C. A.; DA SILVA, S. R. P. Método para Avaliação de Folksonomias como Conceitualizações Compartilhadas na Especificação Colaborativa de Modelos Conceituais. Proceedings of the X Brazilian Symposium in Collaborative Systems, 2013, Sociedade Brasileira de Computação. p.64.

FIALHO, F. A. P. **Gestão do conhecimento e aprendizagem: as estratégias competitivas da sociedade pós-industrial**. Visual Books, 2006. ISBN 8575021923.

FLINT, D. J. Compressing new product success-to-success cycle time: deep customer value understanding and idea generation. **Industrial Marketing Management**, v. 31, n. 4, p. 305-315, 2002. ISSN 0019-8501.

GAUSEMEIER, J.; BERGER, T. Ideas management - Identifying future business ideas. **Konstruktion**, n. 9, p. 64-68, 2004. ISSN 07205953 (ISSN). Disponível em: < <http://www.scopus.com/inward/record.url?eid=2-s2.0-77950426228&partnerID=40&md5=d953a11d30e0fcbc3db67894f7f0f803> >.

GEDIKLI, F. LocalRank – A graph-based tag recommender. In: (Ed.). **Recommender Systems and the Social Web**: Springer Fachmedien Wiesbaden, 2013. cap. 3, p.33-42. ISBN 978-3-658-01947-1.

GELDART, J.; CUMMINS, S. The Automatic Integration of Folksonomies with Taxonomies Using Non-axiomatic Logic. In: PAPADOPOULOS, G. A.; WOJTKOWSKI, W., et al (Ed.). **Information Systems Development**: Springer US, 2010. cap. 38, p.365-372. ISBN 978-0-387-84809-9.

GEMMELL, J. et al. Hybrid tag recommendation for social annotation systems. Proceedings of the 19th ACM international conference on Information and knowledge management, 2010, ACM. p.829-838.

GHOSH, S.; KANE, P.; GANGULY, N. Identifying overlapping communities in folksonomies or tripartite hypergraphs. 20th International Conference Companion on World Wide Web, WWW 2011, 2011, Hyderabad. p.39-40.

GHOSHAL, G. et al. Random hypergraphs and their applications. **Physical Review E - Statistical, Nonlinear, and Soft Matter Physics**, v. 79, n. 6, 2009. ISSN 15393755 (ISSN). Disponível em: < <http://www.scopus.com/inward/record.url?eid=2-s2.0-67650496966&partnerID=40&md5=ea2c9dd13e7808245164b46225dfbb26> >.

GIANGRECO, E. et al. Modeling tools of service value networks to support social innovation in a Smart City. In: JANSSEN, M.; BANNISTER, F., *et al* (Ed.). **Electronic Government and Electronic Participation**, v.21, 2014. p.206-215. (Innovation and the Public Sector). ISBN 978-1-61499-429-9; 978-1-61499-428-2.

GOLDER, S. A.; HUBERMAN, B. A. Usage patterns of collaborative tagging systems. **Journal of Information Science**, v. 32, n. 2, p. 198-208, 2006. ISSN 01655515 (ISSN). Disponível em: < <http://www.scopus.com/inward/record.url?eid=2-s2.0-33644912045&partnerID=40&md5=a9dc5f0d862c65e2a6accdcda5665199> <http://jis.sagepub.com/content/32/2/198.full.pdf> >.

GÓMEZ-PÉREZ, J. M.; ERDMANN, M.; GREAVES, M. **Applying problem solving methods for process knowledge acquisition, representation, and reasoning**. Proceedings of the 4th international conference on Knowledge capture. Whistler, BC, Canada: ACM: 15-22 p. 2007.

GRUBER, T. Collective knowledge systems: Where the Social Web meets the Semantic Web. **Web Semant.**, v. 6, n. 1, p. 4-13, 2008. ISSN 1570-8268.

HEINZLE, R. **Um modelo de engenharia do conhecimento para sistemas de apoio a decisão com recursos para raciocínio abduutivo**. 2011. (Doutorado). Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina, Centro Tecnológico, Florianópolis, SC.

HERSTATT, C. et al. "Fuzzy front end" practices in innovating Japanese companies. **International Journal of Innovation and Technology Management**, v. 3, n. 01, p. 43-60, 2006. ISSN 0219-8770.

HESTER, P. T.; ADAMS, K. M. Systemic thinking: Fundamentals for understanding problems and messes. In: (Ed.). Cham, Switzerland: Springer International, 2014. cap. The Where of Systemic Thinking, p.158. ISBN 978-3-319-07629-4.

HEYMANN, S.; LE GRAND, B. Visual Analysis of Complex Networks for Business Intelligence with Gephi. 2013 17th International Conference on Information Visualisation, 2013, IEEE. p.307-312.

HOLDER, L. B.; MARKOV, Z.; RUSSELL, I. Advances in knowledge acquisition and representation. **International Journal on Artificial Intelligence Tools**, v. 15, n. 06, p. 867-874, 2006. ISSN 0218-2130.

HORTON, G.; GOERS, J. Mining Hidden Profiles in the Collaborative Evaluation of Raw Ideas. System Sciences (HICSS), 2014 47th Hawaii International Conference on, 2014, 6-9 Jan. 2014. p.463-472.

HOTH, A. Data Mining on Folksonomies. In: ARMANO, G.; DE GEMMIS, M., *et al* (Ed.). **Intelligent Information Access**: Springer Berlin Heidelberg, v.301, 2010. cap. 4, p.57-82. (Studies in Computational Intelligence). ISBN 978-3-642-13999-4.

HRASTINSKI, S. et al. A review of technologies for open innovation: Characteristics and future trends. 43rd Annual Hawaii International Conference on System Sciences, HICSS-43, 2010, Koloa, Kauai, HI.

HU, Y. Efficient, high-quality force-directed graph drawing. **Mathematica Journal**, v. 10, n. 1, p. 37-71, 2005. ISSN 1047-5974. Disponível em: <
http://my-canvas-toys.googlecode.com/svn-history/r26/trunk/doc_refs/Alternative_ForceDirected_graph_draw.pdf>.

JÄSCHKE, R. et al. Discovering shared conceptualizations in folksonomies. **Web Semantics: Science, Services and Agents on the World Wide Web**, v. 6, n. 1, p. 38-53, 2008. ISSN 1570-8268. Disponível em: <
<http://www.sciencedirect.com/science/article/pii/S1570826807000546>>.

JELASSI, M. A Quadratic Approach for Trend Detection in Folksonomies. In: KRÖTZSCH, M. e STRACCIA, U. (Ed.). **Web Reasoning and Rule**

Systems: Springer Berlin Heidelberg, v.7497, 2012. cap. 32, p.278-283. (Lecture Notes in Computer Science). ISBN 978-3-642-33202-9.

KANERVA, M.; HOLLANDERS, H. The impact of the economic crisis on innovation. Analysis based on the Innobarometer 2009 survey. **Thematic Paper, European Commission, DG Enterprises, Brussels, dicembre, 2009.**

KARLSSON, A.; TÖRLIND, P. What happens to rejected ideas? - Exploring the life of ideas following the completion of projects. **19th International Conference on Engineering Design, ICED 2013**, Seoul, v. 3 DS75-03, p. 229-238, 2013. ISSN 9781904670469 (ISBN). Disponível em: < <http://www.scopus.com/inward/record.url?eid=2-s2.0-84897670115&partnerID=40&md5=90f1651cf9373ddc09d868c20fee39db> >.

KEMPE, N. et al. An Optimal Algorithm for Raw Idea Selection under Uncertainty. System Science (HICSS), 2012 45th Hawaii International Conference on, 2012, 4-7 Jan. 2012. p.237-246.

KIM, J.; WILEMON, D. Focusing the fuzzy front-end in new product development. **R&D Management**, v. 32, n. 4, p. 269-279, 2002. ISSN 1467-9310.

KOCK, A.; HEISING, W.; GEMÜNDEN, H. G. How Ideation Portfolio Management Influences Front-End Success. **Journal of Product Innovation Management**, v. 32, n. 4, p. 539-555, 2015. ISSN 1540-5885. Disponível em: < <http://dx.doi.org/10.1111/jpim.12217> >.

KOEN, P. et al. Providing Clarity and a Common Language to the **Research-Technology Management**, v. 44, n. 2, p. 46-55, Mar-Apr 2001. ISSN 0895-6308. Disponível em: <<Go to ISI>://WOS:000167289200010 >.

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

KOEN, P. A.; BERTELS, H. M. J.; KLEINSCHMIDT, E. J. Managing the Front End of Innovation—Part I. **Research Technology Management**, Março/Abril 2014a. Disponível em: < <http://www.frontendinnovation.com/media/default/pdfs/fei-article-1.pdf> >.

_____. Managing the Front End of Innovation–Part II. **Research Technology Management**, Maio/Junho 2014b. Disponível em: <
<http://www.frontendinnovation.com/media/default/pdfs/fei-article-2.pdf>>.

KRAUSE, B. et al. Logsonomy - Social information retrieval with logdata. **HYPertext'08: 19th ACM Conference on Hypertext and Hypermedia 2008, HT'08, 2008, Pittsburgh, PA.** p.157-166.

LANDIA, N. et al. Deeper Into the Folksonomy Graph: FolkRank Adaptations and Extensions for Improved Tag Recommendations. **arXiv preprint arXiv:1310.1498**, 2013.

LINDERGAARD, S. A revolução da inovação aberta: a chave da nova competitividade nos negócios. **São Paulo: Évora**, 2011.

LOHMANN, S.; DÍAZ, P. Representing and visualizing folksonomies as graphs - A reference model. 2012 International Working Conference on Advanced Visual Interfaces, AVI 2012, 2012, Capri Island. p.729-732.

MADHUSUDANAN, N.; CHAKRABARTI, A. A questioning based method to automatically acquire expert assembly diagnostic knowledge. **Computer-Aided Design**, v. 57, p. 1-14, 2014. ISSN 0010-4485. Disponível em: <
<http://www.sciencedirect.com/science/article/pii/S0010448514001201>
http://ac.els-cdn.com/S0010448514001201/1-s2.0-S0010448514001201-main.pdf?_tid=3c3ebc38-929e-11e5-8b33-00000aab0f6c&acdnat=1448364581_498b6600f2695b51cd3ee4ac8965c3ec>.

MANUAL DE OSLO. Proposta de diretrizes para coleta e interpretação de dados sobre inovação tecnológica. **OCDE. Organização para Cooperação e Desenvolvimento Econômico: Departamento Estatístico da Comunidade Européia**, 2004.

MARCH, S. T.; SMITH, G. F. Design and natural science research on information technology. **Decision support systems**, v. 15, n. 4, p. 251-266, 1995. ISSN 0167-9236.

MENDES, K. D. S.; SILVEIRA, R. C. D. C. P.; GALVÃO, C. M. Revisão integrativa: método de pesquisa para a incorporação de evidências na saúde e na enfermagem. **Texto and Contexto Enfermagem**, v. 17, n. 4, p. 758, 2008. ISSN 0104-0707.

MIGUEZ, V. B. **Uma abordagem de geração de ideias para o processo de inovação**. 2012. Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina, Centro Tecnológico, Florianópolis, SC.

MIKA, P. Ontologies are us: A unified model of social networks and semantics. **Web semantics: science, services and agents on the World Wide Web**, v. 5, n. 1, p. 5-15, 2007. ISSN 1570-8268. Disponível em: < http://ac.els-cdn.com/S1570826806000552/1-s2.0-S1570826806000552-main.pdf?_tid=14cc389e-2ff0-11e5-be8f-00000aab0f01&acdnat=1437514569_482c34e5f2cc0472c811655a1038c60d >.

MIKELSONE, E.; LIELĀ, E. DISCUSSION ON THE TERMS OF IDEA MANAGAMENT AND IDEA MANAGAMENT SYSTEMS. **DISKUSIJA DĒL IDĒJU VALDYMO IR IDĒJU VALDYMO SISTEMU TERMIŅU.**, n. 17, p. 97-111, 2015. ISSN 20299370. Disponível em: < <http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=110359250&lang=pt-br&site=eds-live&authtype=cookie,ip,uid> >.

MOREIRO GONZÁLEZ, J. A. Papel de los vocabularios semánticos en la economía en red. 12º CONTECSI-FEA-USP e 1º Congresso Internacional em Tecnologia e Organização da Informação, 2015, São Paulo, SP. 22 de maio de 2015. p.9.

MURAH, M. Z. et al. Kacang cerdik: A conceptual design of an idea management system. **International Education Studies**, v. 6, n. 6, p. 178-184, 2013. ISSN 19139020 (ISSN). Disponível em: < <http://www.scopus.com/inward/record.url?eid=2-s2.0-84878690278&partnerID=40&md5=3a63ff3eddbdc45d9c63844f1c4ee6f1>
<http://www.ccsenet.org/journal/index.php/ies/article/download/27887/16836> >.

PACHECO, R. C. D. S. **Dados e Governo Abertos na Sociedade do Conhecimento**. Linked Open Data - Brasil. Florianópolis - SC 2014.

PAGE, L. et al. **The PageRank Citation Ranking: Bringing Order to the Web**. Stanford InfoLab. 1999

PAUKKERI, M. S.; KOTRO, T. Framework for analyzing and clustering short message database of ideas. 9th International Conference on Knowledge Management and Knowledge Technologies, I-KNOW 2009 and

5th International Conference on Semantic Systems, I-SEMANTICS 2009, 2009, Graz. p.239-247.

PEFFERS, K. et al. A design science research methodology for information systems research. **Journal of management information systems**, v. 24, n. 3, p. 45-77, 2007. ISSN 0742-1222.

PENELA, V. et al. miKrow: Semantic Intra-enterprise Micro-Knowledge Management System. In: ANTONIOU, G.; GROBELNIK, M., *et al* (Ed.). **The Semantic Web: Research and Applications**: Springer Berlin Heidelberg, v.6644, 2011. cap. 11, p.154-168. (Lecture Notes in Computer Science). ISBN 978-3-642-21063-1.

PEREZ, A.; LARRINAGA, F.; CURRY, E. **The role of linked data and semantic-technologies for sustainability idea management**. 11th International Conference on Software Engineering and Formal Methods, SEFM 2013 - Collocated Workshops: BEAT2, WS-FMDS, FM-RAIL-Bok, MoKMaSD, and OpenCert. Madrid: Springer Verlag. 8368 LNCS: 306-312 p. 2014.

POMPEO, D. A.; ROSSI, L. A.; GALVÃO, C. M. Revisão integrativa: etapa inicial do processo de validação de diagnóstico de enfermagem. **Acta paul enferm**, v. 22, n. 4, p. 434-8, 2009.

POVEDA, G.; WESTERSKI, A.; IGLESIAS, C. A. Application of semantic search in idea management systems. 7th International Conference for Internet Technology and Secured Transactions, ICITST 2012, 2012, London. p.230-236.

PRADA, C. A. **Proposta de modelo para o gerenciamento de portfólio de inovação: modelagem do conhecimento na geração de ideias**. 2009. 161 (Mestrado). Programa de Pós-graduação em Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina, Centro Tecnológico Florianópolis.

QUINN, A.; WOEHLE, R.; TIEMANN, K. Social network analysis for analyzing groups as complex systems. **Journal of Social Service Research**, v. 38, n. 5, p. 605-618, 2012. ISSN 0148-8376.

REPKO, A. F. **Interdisciplinary Research: Process and Theory**. SAGE Publications, 2011. ISBN 9781412988773. Disponível em: < <https://books.google.com.br/books?id=IOPiSIgmp38C> >.

RIBEIRO JUNIOR, D. I. **Modelo de sistema baseado em conhecimento para apoiar processos de tomada de decisão em ciência e tecnologia.**

2011. Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina, Centro Tecnológico, Florianópolis, SC.

RICCI, F.; ROKACH, L.; SHAPIRA, B. **Introduction to recommender systems handbook.** Springer, 2011. ISBN 0387858199.

RICE, M. P.; O'CONNOR, G. C.; PIERANTOZZI, R. Counter Project Uncertainty. **MIT Sloan Management Review**, 2008.

RICHARDSON, M.; DOMINGOS, P. Markov logic networks. **Machine learning**, v. 62, n. 1-2, p. 107-136, 2006. ISSN 0885-6125.

RICHTER, A.; VOGEL, G. Skibaserl - Knowledge management in high-performance sports. 9th International Conference on Knowledge Management and Knowledge Technologies, I-KNOW 2009 and 5th International Conference on Semantic Systems, I-SEMANTICS 2009, 2009, Graz. p.342-347.

RIEL, A.; NEUMANN, M.; TICHKIEWITCH, S. Structuring the early fuzzy front-end to manage ideation for new product development. **CIRP Annals - Manufacturing Technology**, v. 62, n. 1, p. 107-110, // 2013. ISSN 0007-8506. Disponível em: <
<http://www.sciencedirect.com/science/article/pii/S0007850613001297>>.

RUSSELL, R. K.; TIPPETT, D. D. Critical success factors for the fuzzy front end of innovation in the medical device industry. **Engineering management journal**, v. 20, n. 3, p. 36-43, 2008. ISSN 1042-9247.

SAATY, T. What is the Analytic Hierarchy Process? In: MITRA, G.; GREENBERG, H., *et al* (Ed.). **Mathematical Models for Decision Support**: Springer Berlin Heidelberg, v.48, 1988. cap. 5, p.109-121. (NATO ASI Series). ISBN 978-3-642-83557-5.

SANDSTRÖM, C.; BJÖRK, J. Idea management systems for a changing innovation landscape. **International Journal of Product Development**, v. 11, n. 3-4, p. 310-324, 2010. ISSN 14779056 (ISSN). Disponível em: <
<http://www.scopus.com/inward/record.url?eid=2-s2.0-77954417947&partnerID=40&md5=6fba6f187490cdef582e1c9f2cd2415c>>.

SARIYÜCE, A. E. et al. Streaming algorithms for k-core decomposition. **Proceedings of the VLDB Endowment**, v. 6, n. 6, p. 433-444, 2013. ISSN 2150-8097.

SCHMITT, M. T. B. **Inteligência Competitiva na Web: Um Framework Conceitual para Aquisição de Ativos de Conhecimento no Contexto do Front-End da Inovação**. 2013. 235 (Mestrado). Programa de Pós-graduação em Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina, Florianópolis.

SCHREIBER, G. **Knowledge engineering and management: the CommonKADS methodology**. MIT press, 2000. ISBN 0262193000.

SELART, M.; JOHANSEN, S. T. Understanding the Role of Value-Focused Thinking in Idea Management. **Creativity and Innovation Management**, v. 20, n. 3, p. 196-206, Sep 2011. ISSN 0963-1690. Disponível em: <<Go to ISI>://WOS:000300213900006 >.

SIMON, H. A. **The sciences of the artificial**. MIT press, 1996. ISBN 0262264498.

SMITH, B.; WELTY, C. Ontology: Towards a new synthesis. Formal Ontology in Information Systems, 2001, ACM Press, USA, pp. iii-x. p.3-9.

SPENCER, R. W. The size and shape of "idea space". **International Journal of Innovation Science**, v. 4, n. 2, p. 71-75, 2012. ISSN 17572223 (ISSN). Disponível em: < <http://www.scopus.com/inward/record.url?eid=2-s2.0-84863461876&partnerID=40&md5=ac9d5b7f6a057abdd680efda253bf417> >.

STROHMAIER, M. et al. Evaluation of folksonomy induction algorithms. **ACM Transactions on Intelligent Systems and Technology**, v. 3, n. 4, 2012. ISSN 21576904 (ISSN). Disponível em: < <http://www.scopus.com/inward/record.url?eid=2-s2.0-84864066772&partnerID=40&md5=2aa0a307ef3594450f10b04c2e473cc2> >.

STUDER, R.; BENJAMINS, V. R.; FENSEL, D. Knowledge engineering: Principles and methods. **Data & Knowledge Engineering**, v. 25, n. 1-2, p. 161-197, 3// 1998. ISSN 0169-023X. Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0169023X97000566> >.

SU, X.; KHOSHGOFTAAR, T. M. A survey of collaborative filtering techniques. **Adv. in Artif. Intell.**, v. 2009, p. 2-2, 2009. ISSN 1687-7470.

TAGLINO, F.; SMITH, F.; PROIETTI, M. Knowledge-based support to business innovation. CEUR Workshop Proceedings, 2012. p.37-44.

TANWAR, P.; PRASAD, T.; ASWAL, M. S. Comparative study of three declarative knowledge representation techniques. **International Journal on Computer Science and Engineering**, v. 2, n. 07, p. 2274-2281, 2010.

TEZA, P. **Front end da inovação: proposta de um modelo conceitual**. 2012. 147 (Mestrado). Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis.

TIDD, J.; BESSANT, J. **Gestão da inovação**. 5. Porto Alegre, RS: Bookman Editora, 2015. ISBN 858260307X.

TSUI, E. et al. A concept–relationship acquisition and inference approach for hierarchical taxonomy construction from tags. **Information Processing & Management**, v. 46, n. 1, p. 44-57, 1// 2010. ISSN 0306-4573. Disponível em: <
<http://www.sciencedirect.com/science/article/pii/S030645730900065X>
http://ac.els-cdn.com/S030645730900065X/1-s2.0-S030645730900065X-main.pdf?_tid=fd5ec1b0-953c-11e5-a4c0-00000aacb35f&acdnat=1448652668_38687241de5e45f7813c24f27db234a1>.

TUDORACHE, T. et al. WebProtégé: A collaborative ontology editor and knowledge acquisition tool for the Web. **Semantic Web**, v. 4, n. 1, p. 89, 2013.

VAN AKEN, J. E.; ROMME, A. G. L. A design science approach to evidence-based management. **The Oxford handbook of evidence-based management**, p. 43-61, 2012.

VAN ELST, L.; ABECKER, A. Ontologies for information management: balancing formality, stability, and sharing scope. **Expert Systems with Applications**, v. 23, n. 4, p. 357-366, 11// 2002. ISSN 0957-4174. Disponível em: <
<http://www.sciencedirect.com/science/article/pii/S0957417402000714>>.

VANDENBOSCH, B.; SAATCIOGLU, A.; FAY, S. Idea Management: A Systemic View*. **Journal of Management Studies**, v. 43, n. 2, p. 259-288,

2006. ISSN 1467-6486. Disponível em: < <http://dx.doi.org/10.1111/j.1467-6486.2006.00590.x> >.

VON ALAN, R. H. et al. Design science in information systems research. **MIS quarterly**, v. 28, n. 1, p. 75-105, 2004.

WESTERSKI, A. **Semantic technologies in idea management systems: a model for interoperability, linking and filtering**. Adam Westerski, 2013.

WESTERSKI, A.; DALAMAGAS, T.; IGLESIAS, C. A. Classifying and comparing community innovation in Idea Management Systems. **Decision Support Systems**, v. 54, n. 3, p. 1316-1326, 2013. ISSN 01679236 (ISSN). Disponível em: < <http://www.scopus.com/inward/record.url?eid=2-s2.0-84875421321&partnerID=40&md5=a3f5b8809e66d638855b8b01efcc47f0> >.

WESTERSKI, A.; IGLESIAS, C. A. Exploiting structured linked data in enterprise knowledge management systems: An idea management case study. 15th IEEE International EDOC Enterprise Computing Conference Workshops, EDOCW 2011, 2011, Helsinki. p.395-403.

WESTERSKI, A.; IGLESIAS, C. A.; GARCIA, J. E. Idea relationship analysis in open innovation crowdsourcing systems. 8th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing, CollaborateCom 2012, 2012, Pittsburgh, PA. p.289-296.

WESTERSKI, A.; IGLESIAS, C. A.; NAGLE, T. The road from community ideas to organisational innovation: A life cycle survey of idea management systems. **International Journal of Web Based Communities**, v. 7, n. 4, p. 493-506, 2011. ISSN 14778394 (ISSN). Disponível em: < <http://www.scopus.com/inward/record.url?eid=2-s2.0-80053949856&partnerID=40&md5=861265f82a1afdc69cca092776350f7c> >.

WESTERSKI, A.; IGLESIAS, C. A.; RICO, F. T. **A model for integration and interlinking of idea management systems**. 4th International Conference on Metadata and Semantic Research, MTSR 2010. Alcalá de Henares. 108 CCIS: 183-194 p. 2010.

WILBERT, J. K. W. **Características de VCoPs que Influenciam Processos de Inovação: Estudo de Caso em uma Empresa Pública Brasileira**. 2015. (Mestrado). Programa de Pós-Graduação em Engenharia

e Gestão do Conhecimento, Universidade Federal de Santa Catarina, Centro Tecnológico, Florianópolis, SC.

XU, K. et al. A Comparative Study of Correlation Measurements for Searching Similar Tags. In: TANG, C.;LING, C., *et al* (Ed.). **Advanced Data Mining and Applications**: Springer Berlin Heidelberg, v.5139, 2008. cap. 75, p.709-716. (Lecture Notes in Computer Science). ISBN 978-3-540-88191-9.

ZHANG, Q.; ZHANG, Q.; PENG, X. Research on knowledge service system in open innovation environment. International Conference on Management and Service Science, MASS 2011, 2011.

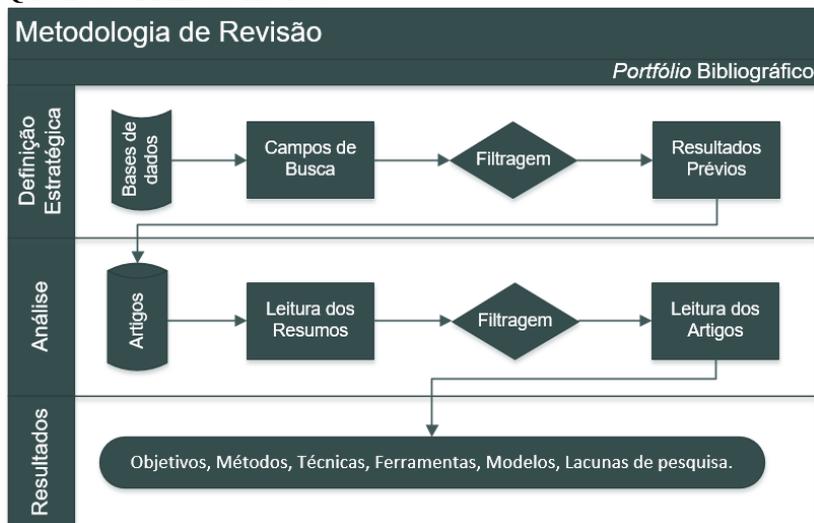
ZHANG, Z. K.; LIU, C. A hypergraph model of social tagging networks. **Journal of Statistical Mechanics: Theory and Experiment**, v. 2010, n. 10, 2010. ISSN 17425468 (ISSN). Disponível em: <
<http://www.scopus.com/inward/record.url?eid=2-s2.0-78650397530&partnerID=40&md5=3814ad5b0f889479e85c8e81e8da38be>
>.

APÊNDICE A – METODOLOGIA DE REVISÃO

As estratégias de busca e processos descritos apresentam uma revisão bibliográfica integrativa, de caráter exploratório, em uma abordagem qualitativa por apresentar uma revisão sistemática sobre os objetivos, ferramentas e métodos, e por ter a finalidade de realizar uma análise do conhecimento pré-existente sobre os tópicos pesquisados (Pompeo *et al.*, 2009). A presente seção apresenta os procedimentos de pesquisa realizados para a seleção e análise do material bibliográfico selecionado.

Nas seções seguintes seguem as etapas apresentadas por Mendes *et al.* (2008), que integram a definição do tema (elaboração da Pergunta de Pesquisa); a busca na literatura (amostragem da base de dados online); os critérios para categorização dos estudos (filtragem de dados); a avaliação dos estudos incluídos; a discussão do resultado e a apresentação da revisão integrativa. O fluxo dessas etapas da pesquisa pode ser conferido no.

Quadro 5 - Fluxo da Revisão



Fonte: do autor.

O propósito desta revisão integrativa é compreender o estado da arte relacionado aos Sistemas de Gestão de Ideias e outra pesquisa relacionada aos estudos das Folksonomias.

APÊNDICE B – BIBLIOMETRIA DE IMS

A partir dos objetivos e pergunta chave, foram definidos como palavras de busca os termos em inglês “*idea management system*” ou, no plural, “*idea management systems*”, realizando as devidas filtrações para abordagem do tema e objetivos estabelecidos.

As definições das estratégias de busca foram estabelecidas nos campos de busca, filtração e resultados prévios, conforme segue:

- 1- **Campos de busca:** A pesquisa foi realizada dentro das principais bases de dados. Os campos de busca foram o título do artigo, resumo e palavras-chave. A filtração realizada na base SCOPUS foi: “(TITLE-ABS-KEY (“**idea management system**”) OR TITLE-ABS-KEY (“**idea management systems**”))”, obtendo 38 documentos. Em seguida na base Web of Science (WoS): “**Tópico: (“idea management systems”)** OR **Tópico: (“idea management systems”)** ”, obtendo 9 documentos.
- 2- **Filtração:** Os resultados retornaram alguns resumos de conferências que não abordavam o tema, apenas possuíam estas palavras-chave. Assim, foram obtidos 36 artigos na base SCOPUS e 9 na base WoS.
- 3- **Resultados prévios:** Dos 45 documentos, excluindo repetidos, apenas dois não estavam disponíveis na base WoS (Selart e Johansen, 2011; Giangreco *et al.*, 2014), assim, a base manteve 38 artigos, apresentados no Quadro 10.

A busca foi realizada em fevereiro de 2015 e atualizada em janeiro de 2016. Do portfólio bibliográfico selecionado, 19 documentos são artigos de conferências, 16 artigos em revistas e três revisões.

As principais áreas do conhecimento são: Ciência da Computação (16 artigos); Engenharia (13 artigos); Negócios, Gestão e Contabilidade (oito artigos), a lista completa de documentos por área é apresentada no Quadro 6.

O Quadro 7 apresenta as palavras-chave com maior ocorrência no conjunto de artigos selecionados. Foram agrupados os termos apresentados em plural e singular somando o número de ocorrências. Dentre as palavras comuns relacionadas como “Inovação”, “Gestão” e “Ideias” estão os termos “Gestão do Conhecimento” e “Inteligência Coletiva”, com oito e seis publicações, respectivamente.

Quadro 6 – Documentos por área do conhecimento.

ÁREAS DO CONHECIMENTO	QNT
Computer Science	16
Engineering	13
Business, Management and Accounting	8
Mathematics	5
Decision Sciences	4
Economics, Econometrics and Finance	2
Social Sciences	2
Outros (Arts and Humanities, Chemical Engineering, Materials Science, Psychology)	6

Fonte: do autor.

Quadro 7 – Palavras-chave por ocorrência.

PALAVRAS-CHAVE	QNT
Idea management/Idea managements	12
Idea management system/Idea management systems	12
Management	12
Innovation	11
Management systems	10
Idea management systems	8
Knowledge management	8
Collective intelligence/Collective intelligences	6
Design	6
Product development	6

Fonte: do autor.

Outras palavras-chave com mais de duas ocorrências são: Idea generation (4); Information management (4); innovation management (4); Life cycle (4); ontology (4); Open innovation (4); Semantic Web (4); Semantics (4); community (3); Computer supported; cooperative work (3); Costs (3); Industrial management (3); Information system/Information systems (3); Interactive computer systems (3); Metadata (3); Problem solving (3); Process improvement (3); Semantic Web technology (3); Business organizations (2); Competition (2); Cost effectiveness (2); creative idea/Creative ideas (2); creativity (2); Crowdsourcing (2); Decision making/decision; making (2); deliberation (2); Development project (2); Groupware (2); Idea (2); Industry (2); Information systems (2); Innovation process (2); Innovative ideas (2);

Knowledge acquisition (2); Linked data (2); Linked datum (2); Management innovation (2); Management strategies (2) e Price setting (2); Social sciences (2). A Figura 30 apresenta uma nuvem de *tag* criadas a partir das coocorrências.

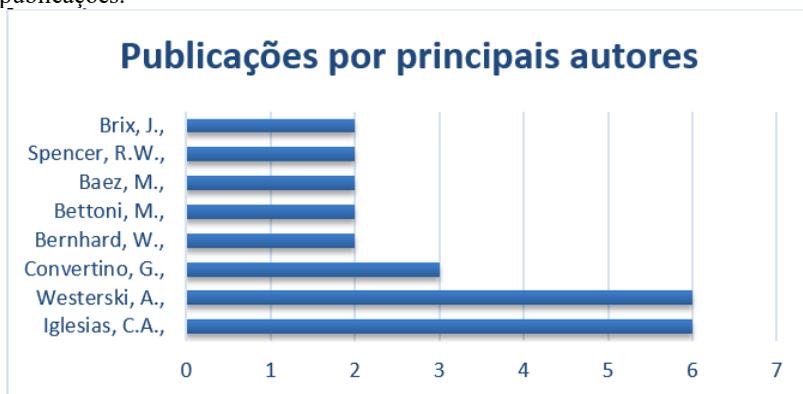
Figura 30 – Nuvem de *Tags* a partir da busca sobre Sistemas de Gestão de Ideias (IMSS).



Fonte: do autor.

O Gráfico 2, apresenta o número de publicações por autor. Os principais autores relacionados ao tema de IMS são C.A. Iglesias e A. Westerski da Universidad Politecnica de Madrid, com seis publicações em conjunto. Em especial citando o sistema Gi2MO desenvolvido pelo grupo (Westerski *et al.*, 2013).

Gráfico 2 – Número de publicações por autores, com no mínimo duas publicações.



Fonte: do autor.

Quadro 8 – Publicações por instituição.

INSTITUIÇÃO	QNT
Universidad Politecnica de Madrid	6
Xerox Research Centre Europe	3
Palo Alto Research Center	3
Chalmers University of Technology	2
Universita degli Studi di Trento	2

Fonte: do autor.

Esses autores estão ligados as instituições apresentadas no Quadro 8. Entre as que mais publicaram, além da Universidad Politecnica de Madrid C.A. Iglesias e A. Westerski, estão a “Xerox Research Centre Europe” e “Palo Alto Research Center Chalmers University of Technology”.

O Gráfico 3 apresenta o total de publicações por país, considerando os que apresentam dois ou mais documento. O grande destaques é Espanha e Estados Unidos com sete publicações cada.

Gráfico 3 – Publicações por país.



Fonte: do autor.

Gráfico 4 – Artigos por ano de publicação.



Fonte: do autor.

O Gráfico 4 apresenta o número de publicações por ano. É possível perceber o aumento de publicações sobre o assunto após a crise de 2008, mantendo uma quantidade semelhante de publicações nos últimos anos, porém ainda baixo.

O Quadro 9 apresenta as fontes das publicações sobre IMS com mais de uma publicação. Sendo a “Proceedings of the ACM Conference on Computer Supported Cooperative Work CSCW” a principal fonte de publicações sobre o assunto.

Por fim, o Quadro 10 mostra um mapeamento das publicações apresentadas na revisão sistemática e o número de citações nas bases Scopus e Web of Science (WoS). Publicações não encontradas em alguma dessas bases foram identificadas por “ND” na respectiva coluna, o que significa: “Não definido”. Os artigos estão ordenados por ano, a fim de apresentar quais as novidades, tendências e gaps na área.

Quadro 9 – Fontes das publicações.

JOURNAL	QNT
Proceedings of the ACM Conference on Computer Supported Cooperative Work CSCW	3
Applied Mechanics and Materials	2
International Journal of Product Development	2

Fonte: do autor.

Quadro 10 – Mapeamento dos Artigos de IMS e Dados Bibliométricos.

ANO	TÍTULO	AUTORES	TÍTULO DO JOURNAL	SCOPUS	WoS
2015	The performance-improving benefits of a radical innovation initiative	Brix J.	International Journal of Productivity and Performance Management	1	ND
2015	Applying Idea Management System (IMS) approach to design and implement a collaborative environment in public service related open innovation processes	Alessi M., Camilo A., Chetta V., Giangreco E., Soufivand M., Storelli D.	CEUR Workshop Proceedings	0	ND
2015	IT-enabled product innovation: Customer motivation for participating in virtual idea communities	Bretschneider U., Leimeister J.M., Mathiassen L.	International Journal of Product Development	0	ND
2014	The study on small enterprise innovation	Ling L.M., Liu X.P.	Applied Mechanics and Materials	0	ND
2014	The study on promotion of Chinese green energy-saving buildings	Li W.F.	Applied Mechanics and Materials	0	ND
2014	Relationship between the features as well as industrial organization form of creative idea and creative idea management system	Shi Q.C., Tuo X.N.	Proceedings - 2014 7th International Joint Conference on Computational Sciences and Optimization, CSO 2014	0	ND
2014	The role of linked data and semantic-technologies for sustainability idea management	Perez A., Larrinaga F., Curry E.	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes	0	0

ANO	TÍTULO	AUTORES	TÍTULO DO JOURNAL	SCOPUS	WoS
			in Bioinformatics)		
2014	Y Beyond the Generation of Ideas: Virtual Idea Campaigns to Spur Creativity and Innovation	Elerud-Tryde, Anne; Hooge, Sophie	CREATIVITY AND INNOVATION MANAGEMENT	1	2
2014	Modelling tools of service value networks to support social innovation in a Smart City	Giangreco, Enza; Marasso, Lanfranco; Chetta, Valentina; Fortunato, Laura; Perlangeli, Cosimo	ELECTRONIC GOVERNMENT AND ELECTRONIC PARTICIPATION	ND	0
2013	What happens to rejected ideas? - Exploring the life of ideas following the completion of projects	Karlsson A., Torlind P.	Proceedings of the International Conference on Engineering Design, ICED	0	ND
2013	Collaborative solutions quick&clean: The SFM method	Bettoni M., Bernhard W., Bittel N.	Proceedings of the European Conference on Knowledge Management, ECKM	0	ND
2013	Kacang cerdas: A conceptual design of an idea management system	Murah M.Z., Abdullah Z., Hassan R., Bakar M.A., Mohamed I., Amin H.M.	International Education Studies	1	ND
2013	The establishment of prophase innovative management system in nuclear power project	Wu Y., Naren M., Wang Q., Wang H., Chen J.	Advanced Materials Research	0	ND
2013	Classifying and comparing community innovation in Idea Management Systems	Westerski A., Dalamagas T., Iglesias C.A.	Decision Support Systems	4	4

ANO	TÍTULO	AUTORES	TÍTULO DO JOURNAL	SCOPUS	WoS
2012	Idea relationship analysis in open innovation crowdsourcing systems	Westerski A., Iglesias C.A., Garcia J.E.	CollaborateCom 2012 - Proceedings of the 8th International Conference on Collaborative Computing: Networking, Applications and Worksharing	0	ND
2012	Application of semantic search in idea management systems	Poveda G., Westerski A., Iglesias C.A.	2012 International Conference for Internet Technology and Secured Transactions, ICITST 2012	0	0
2012	The size and shape of "idea space"	Spencer R.W.	International Journal of Innovation Science	1	ND
2012	Collective intelligence as community discourse and action	De Liddo A., Buckingham Shum S., Convertino G., Sandor A., Klein M.	Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW	0	ND
2012	Designing a facilitator's cockpit for an idea management system	Baez M., Convertino G.	Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW	1	ND
2012	Innovation cockpit: A dashboard for facilitators in idea management	Baez M., Convertino G.	Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW	0	ND
2011	Exploiting structured linked data in	Westerski A., Iglesias	Proceedings - IEEE International	4	ND

ANO	TÍTULO	AUTORES	TÍTULO DO JOURNAL	SCOPUS	WoS
	enterprise knowledge management systems: An idea management case study	C.A.	Enterprise Distributed Object Computing Workshop, EDOC		
2011	The road from community ideas to organisational innovation: A life cycle survey of idea management systems	Westerski A., Iglesias C.A., Nagle T.	International Journal of Web Based Communities	9	ND
2011	Understanding the Role of Value-Focused Thinking in Idea Management	Selart, Marcus; Johansen, Svein Tvedt	CREATIVITY AND INNOVATION MANAGEMENT	ND	4
2010	A model for integration and interlinking of idea management systems	Westerski A., Iglesias C.A., Rico F.T.	Communications in Computer and Information Science	2	1
2010	Idea management by role based networked learning	Bettoni M., Bernhard W., Eggs C., Schiller G.	Proceedings of the European Conference on Knowledge Management, ECKM	2	ND
2010	Idea management system for team creation	Xie L., Zhang P.	Journal of Software	2	ND
2010	What's your idea? A case study of a grassroots innovation pipeline within a large software company	Bailey B.P., Horvitz E.	Conference on Human Factors in Computing Systems - Proceedings	6	2
2010	Idea management systems for a changing innovation landscape	Sandstrom C., Bjork J.	International Journal of Product Development	9	ND
2010	A review of technologies for open innovation: Characteristics and future trends	Hrastinski S., Kviselius N.Z., Ozan H., Edenius M.	Proceedings of the Annual Hawaii International Conference on System Sciences	0	ND

ANO	TÍTULO	AUTORES	TÍTULO DO JOURNAL	SCOPUS	WoS
2009	Skibaserl - Knowledge management in high-performance sports	Richter A., Vogel G.	Proceedings of I-KNOW 2009 - 9th International Conference on Knowledge Management and Knowledge Technologies and Proceedings of I-SEMANTICS 2009 - 5th International Conference on Semantic Systems	0	ND
2007	Innovation by the side door	Spencer R.W.	Research Technology Management	3	ND
2006	Supporting cost saving ideas reuse with an ontology based tool	Angeniol S., Yannou B., Longueville B., Chamerois R., Gardoni M.	Proceedings of the ASME Design Engineering Technical Conference	0	ND
2004	Ideas management - Identifying future business ideas Ideenmanagement in der strategischen produktplanung - Identifikation der produkte und geschäftsfelder von morgen	Gausemeier J., Berger T.	Konstruktion	0	ND
2004	Idea, management system and planning method of urban fringe development	Li H.-P., Li J.-L.	Chongqing Jianzhu Daxue Xuebao/Journal of Chongqing Jianzhu University	1	ND
2003	Inventoried oil saves millions		Saudi Aramco Journal of Technology	0	ND

ANO	TÍTULO	AUTORES	TÍTULO DO JOURNAL	SCOPUS	WoS
2002	Lines of code	Chanterlayne I.	Engineering (London)	0	ND
2002	Managing ideas for the development of new products	Nilsson L., Elg M., Bergman B.	International Journal of Technology Management	5	3
1997	Turning technical groups into high-performance teams	Grossman Stephen	IEEE Engineering Management Review	1	ND

Fonte: do autor.

APÊNDICE C – BIBLIOMETRIA DE FOLKSONOMIAS

A partir dos objetivos e pergunta chave, foram definidos como palavras de busca os termos em inglês “*folksonomy*” ou, no plural, “*folksonomies*”, e “*graph*”, realizando as devidas filtrações para abordagem do tema e objetivos estabelecidos.

As definições das estratégias de busca foram estabelecidas nos campos de busca, filtração e resultados prévios, conforme segue:

- 1- **Campos de busca:** A pesquisa foi realizada dentro da base de dados SCOPUS. Os campos de busca foram o título do artigo, resumo e palavras-chave. A filtração realizada foi: “**((TITLE-ABS-KEY (folksonomy) OR TITLE-ABS-KEY (folksonomies)) AND TITLE-ABS-KEY (graph))**”, obtendo 116 documentos.
- 2- **Filtração:** Os resultados retornaram alguns resumos de conferências que não abordavam o tema, apenas possuíam estas palavras-chave. Assim, foram obtidos 99 artigos.
- 3- **Resultados prévios:** Os idiomas dos documentos são predominantemente o inglês, com 97 artigos, em chinês e francês apenas um cada. As publicações estão entre 2007 e 2016.

A busca foi realizada em junho de 2015 e atualizada em janeiro de 2016. Do portfólio bibliográfico selecionado, 67 documentos são artigos de conferências e 26 artigos em Journal, capítulos de livros, artigos impressos e revisões possuem dois documentos cada.

As principais áreas do conhecimento são: Ciência da Computação (86 artigos); Matemática (26 artigos) e Engenharia (nove artigos), a lista completa de documentos por área é apresentada no Quadro 11.

O Quadro 12 apresenta as palavras-chave com maior ocorrência no conjunto de artigos selecionados. Foram agrupados os termos apresentados em plural e singular somando o número de ocorrências. Além do termo chave “*folksonomy*”, os termos que mais se destacam na pesquisa estão relacionados a “semântica” e “redes sociais”. Outras palavras que se destacaram estão relacionadas à análise das folksonomias como: classificação, recomendação, agrupamento e as derivações de hipergrafo como tripartido e bipartido.

Quadro 11 – Documentos por área do conhecimento.

ÁREAS DO CONHECIMENTO	QNT
Computer Science	86
Mathematics	26
Engineering	9
Decision Sciences	6
Physics and Astronomy	6
Social Sciences	6
Business, Management and Accounting	5
Biochemistry, Genetics and Molecular Biology	4
Neuroscience	2
Outros (Agricultural and Biological Sciences, Arts and Humanities, Earth and Planetary Sciences e Medicine)	4

Fonte: do autor.

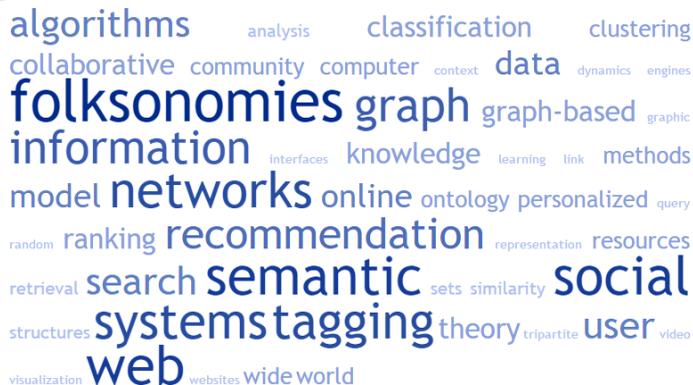
Quadro 12 – Palavras-chave por ocorrência.

PALAVRAS-CHAVE	QNT
Folksonomies/folksonomy	105
Semantic Web/semantic Webs	35
Semantics	24
Graph theory	18
Social networking (online)	18
Social networks/Social Network	17
Tag recommendation/Tag recommendations	16
World Wide Web	16
Classification (of information)	15
Algorithms	14
Data sets	11
User interfaces	11
Graph-based	10
Graphic methods	10
Hyper graph/Hyper graphs/Hypergraph	9
Social tagging	9
Websites	9

Fonte: do autor.

Outras palavras-chave com mais de duas ocorrências foram: Collaborative tagging (8); Information retrieval (8); ranking (8); recommender systems (8); search (8); Ontology (7); Tagging (7); Web resources (7); Hypertext systems (6); personalization/Personalizations (6); Population Dynamics (6); Search engine/Search engines (6); social bookmarking (6); social media (6); systems (6); tripartite graph/Tripartite graphs (6); community detection (5); Information systems (5); Knowledge management (5); Random Walk/random; walk (5); Web 2.0 (5); Bipartite graphs (4); Clustering algorithms (4); Collaborative systems (4); Context (4); FolkRank (4); Forecasting (4); Humans (4); Information theory (4); Knowledge representation (4); Link prediction (4); Multi agent systems (4); Online systems (4); personalized search (4); Classification (3); collaborative filtering (3); complex networks (3); Co-occurrence (3); dominating set/Dominating sets (3); Electric network analysis (3); Graph model (3); Graph; based methods (3); graph; based ranking (3); International conferences (3); model/Models (3); Models, Biological (3); Navigation (3); personalized recommendation (3); Semantic similarity (3); Social Networking (3); social tagging systems (3); *tag* similarity/*tag* similarities (3); Tagging systems (3); Thesauri (3); Topic model/Topic models (3); Tripartite structures (3); User experience (3); User profile/user profiles (3); Visualization (3) e Web users (3). A Figura 31 – Nuvem de *tag* a partir da pesquisa na área de Folksonomias e grafos. apresenta uma nuvem de *tag* criadas a partir das coocorrências

Figura 31 – Nuvem de *tag* a partir da pesquisa na área de Folksonomias e grafos.



Fonte: do autor.

Gráfico 5 – Número de publicações dos principais autores.



Fonte: do autor.

Ao todo são 157 autores com publicações sobre o assunto. O Gráfico 5 apresenta os autores com cinco ou mais publicações.

Esses autores estão ligados as instituições apresentadas no Quadro 13. As principais instituições são da Universitat Kassel da Alemanha e University of Southampton do Reino Unido. Além dessas, outras 106 instituições apresentam publicações sobre o tema, o que apresenta uma grande colaboração.

Quadro 13 – Publicações por instituição.

INSTITUIÇÃO	QNT
Universitat Kassel	6
University of Southampton	6
Universitat Hannover	5
University of Ottawa Canada	5
Universitat Hildesheim	4
City University of Hong Kong	4
Harbin Institute of Technology	4
Indiana University	4

Fonte: do autor.

A grande colaboração entre instituições se reflete na quantidade de países, são 36 diferentes países com publicações na área. O Gráfico 6 apresenta o total de publicações por país, considerando os países que tiveram cinco ou mais documentos. O grande destaque é a Alemanha com 25 documentos, China com 18 e Estados Unidos com 15.

Gráfico 6 – Publicações por país



Fonte: do autor.

O Gráfico 7 confirma o aumento de publicações sobre o assunto após a crise de 2008, mantendo uma quantidade semelhante de publicações nos últimos anos, porém com queda a partir de 2012.

Gráfico 7 – Artigos por ano de publicação.



Fonte: do autor.

Quadro 14 – Fontes das publicações.

JOURNAL	QNT
International Conference on Information and Knowledge Management Proceedings	3
Physical Review E Statistical Nonlinear and Soft Matter Physics	3
Journal of Information Science	2
Multimedia Tools and Applications	2

Fonte: do autor.

O Quadro 14 apresenta as fontes das publicações sobre IMS com mais de uma publicação. A principal fonte de artigos sobre folksonomias e grafos é a “International Conference on Information and Knowledge Management Proceedings” e “Physical Review E Statistical Nonlinear and Soft Matter Physics” com três publicações cada.

No Quadro 15 são apresentadas as principais literaturas sobre o tema ordenadas por número de citações, segundo a base Scopus. As demais publicações da pesquisa, omitidas no quadro, possuem menos de duas citações por serem novas ou por não terem se consolidado.

Quadro 15 – Mapeamento dos Artigos de Folksonomias e Dados Bibliométricos com duas ou mais citações.

ANO	TÍTULO	AUTORES	TÍTULO DO JORNAL	SCOPUS
2007	Ontologies are us: A unified model of social networks and semantics	Mika P.	Web Semantics	315
2005	Ontologies are us: A unified model of social networks and semantics	Mika P.	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	130
2008	<i>Tag</i> recommendations in social bookmarking systems	Jaschke R., Marinho L., Hotho A., Schmidt-Thieme L., Stumme G.	AI Communications	115
2010	Personalized recommendation via integrated diffusion on user-item- <i>tag</i> tripartite graphs	Zhang Z.-K., Zhou T., Zhang Y.-C.	Physica A: Statistical Mechanics and its Applications	89
2007	<i>Tag</i> recommendations in folksonomies	Jaschke R., Marinho L., Hotho A., Schmidt-Thieme L., Stumme G.	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	65
2009	Random hypergraphs and their applications	Ghoshal G., Zlatic V., Caldarelli G., Newman M.E.J.	Physical Review E - Statistical, Nonlinear, and Soft Matter Physics	47
2009	Telling experts from spammers: Expertise ranking in folksonomies	Noll M.G., Au Yeung C.-M., Gibbins N., Meinel C., Shadbolt N.	Proceedings - 32nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2009	32

ANO	TÍTULO	AUTORES	TÍTULO DO JORNAL	SCOPUS
2009	Hypergraph topological quantities for tagged social networks	Zlatic V., Ghoshal G., Caldarelli G.	Physical Review E - Statistical, Nonlinear, and Soft Matter Physics	28
2008	Folksonomies and clustering in the collaborative system CiteULike	Capocci A., Caldarelli G.	Journal of Physics A: Mathematical and Theoretical	17
2009	Contextualising <i>tag</i> in collaborative tagging systems	Yeung C.-M.A., Gibbins N., Shadbolt N.	Proceedings of the 20th ACM Conference on Hypertext and Hypermedia, HT'09	14
2006	Cloudalicious: Folksonomy over time	Russell T.	Proceedings of the ACM/IEEE Joint Conference on Digital Libraries	14
2007	<i>Tag</i> meaning disambiguation through analysis of tripartite structure of folksonomies	Yeung C.-M.A., Gibbins N., Shadbolt N.	Proceedings - 2007 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology - Workshops, WI-IAT Workshops 2007	12
2007	<i>Tag</i> recommendations in folksonomies	Jaschke R., Marinho L., Hotho A., Schmidt-Thieme L., Stumme G.	LWA 2007 - Lernen - Wissen - Adaptivitat - Learning, Knowledge, and Adaptivity, Workshop Proceedings	12
2014	Community-aware user profile enrichment in folksonomy	Xie H., Li Q., Mao X., Li X., Cai Y., Rao Y.	Neural Networks	10
2008	Logsonomy - Social information retrieval with logdata	Krause B., Hotho A., Jaschke R., Stumme G.	HYPERTEXT'08: Proceedings of the 19th ACM Conference on Hypertext and Hypermedia, HT'08 with Creating'08 and WebScience'08	10

ANO	TÍTULO	AUTORES	TÍTULO DO JORNAL	SCOPUS
2011	Personalized PageRank vectors for <i>tag</i> recommendations: Inside FolkRank	Kim H.-N., El Saddik A.	RecSys'11 - Proceedings of the 5th ACM Conference on Recommender Systems	9
2011	Spear: Spamming-resistant expertise analysis and ranking in collaborative tagging systems	Yeung C.-M.A., Noll M.G., Gibbins N., Meinel C., Shadbolt N.	Computational Intelligence	9
2010	Semantic social network analysis: A concrete case	Ereteo G., Limpens F., Gandon F., Corby O., Buffa M., Leitzelman M., Sander P.	Handbook of Research on Methods and Techniques for Studying Virtual Communities: Paradigms and Phenomena	9
2009	<i>Tag</i> navigation	Mesnage C.S., Carman M.J.	SoSEA'09 - Proceedings of the 2nd International Workshop on Social Software Engineering and Applications	9
2014	Mining latent user community for <i>tag</i> -Based and content-Based search in social media	Xie H., Li Q., Mao X., Li X., Cai Y., Zheng Q.	Computer Journal	8
2011	Improving graph-based approaches for personalized <i>tag</i> recommendation	Ramezani M.	Journal of Emerging Technologies in Web Intelligence	8
2010	Ontology emergence from folksonomies	Liu K., Fang B., Zhang W.	International Conference on Information and Knowledge Management, Proceedings	8
2008	Ranking in folksonomy systems: Can context help?	Abel F., Henze N., Krause D.	International Conference on Information and Knowledge Management, Proceedings	8

ANO	TÍTULO	AUTORES	TÍTULO DO JORNAL	SCOPUS
2008	A comparison of social bookmarking with traditional search	Krause B., Hotho A., Stumme G.	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	8
2012	Alignment and integration of complex networks by hypergraph-based spectral clustering	Michoel T., Nachtergaele B.	Physical Review E - Statistical, Nonlinear, and Soft Matter Physics	7
2010	The link prediction problem in bipartite networks	Kunegis J., De Luca E.W., Albayrak S.	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	7
2009	Enrichment and ranking of the Youtube <i>tag</i> space and integration with the Linked Data cloud	Choudhury S., Breslin J.G., Passant A.	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	7
2012	Navigational efficiency of broad vs. narrow folksonomies	Helic D., Korner C., Granitzer M., Strohmaier M., Trattner C.	HT'12 - Proceedings of 23rd ACM Conference on Hypertext and Social Media	6
2010	Semantic history map: Graphs aiding Web revisitation support	Simko J., Tvarozek M., Bielikova M.	Proceedings - 21st International Workshop on Database and Expert Systems Applications, DEXA 2010	6
2010	A graph-based clustering scheme for identifying related <i>tag</i> in folksonomies	Papadopoulos S., Kompatsiaris Y., Vakali A.	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	6

ANO	TÍTULO	AUTORES	TÍTULO DO JORNAL	SCOPUS
2009	On the effect of group structures on ranking strategies in folksonomies	Abel F., Henze N., Krause D., Kriesell M.	Weaving Services and People on the World Wide Web	6
2008	Social networking for scientists using tagging and shared bookmarks: A Web 2.0 application	Pierce M.E., Fox G.C., Rosen J., Maini S., Choi J.Y.	2008 International Symposium on Collaborative Technologies and Systems, CTS'08	6
2011	SemTagP: Semantic community detection in folksonomies	Ereteo G., Gandon F., Buffa M.	Proceedings - 2011 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2011	5
2014	On improving aggregate recommendation diversity and novelty in folksonomy-based social systems	Wu H., Cui X., He J., Li B., Pei Y.	Personal and Ubiquitous Computing	4
2011	Identifying overlapping communities in folksonomies or tripartite hypergraphs	Ghosh S., Kane P., Ganguly N.	Proceedings of the 20th International Conference Companion on World Wide Web, WWW 2011	4
2010	Ontology induction based on social annotations	Liu K.-P., Fang B.-X.	Jisuanji Xuebao/Chinese Journal of Computers	4
2008	Exploiting additional context for graph-based tag recommendations in folksonomy systems	Abel F., Henze N., Krause D.	Proceedings - 2008 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2008	4
2014	Adaptive content recommendation for mobile users: Ordering recommendations using a hierarchical context model with granularity	Han J., Schmidtke H.R., Xie X., Woo W.	Pervasive and Mobile Computing	3

ANO	TÍTULO	AUTORES	TÍTULO DO JORNAL	SCOPUS
2013	Exploiting <i>tag</i> similarities to discover synonyms and homonyms in folksonomies	Eynard D., Mazzola L., Dattolo A.	Software - Practice and Experience	3
2012	Ontologies and <i>tag</i> -statistics	Tibely G., Pollner P., Vicsek T., Palla G.	New Journal of Physics	3
2010	Random graph generative model for Folksonomy network structure approximation	Chojnacki S., Klopotek M.	Procedia Computer Science	3
2010	Visit me, click me, be my friend: An analysis of evidence networks of user relationships in BibSonomy	Mitzlaff F., Benz D., Stumme G., Hotho A.	HT'10 - Proceedings of the 21st ACM Conference on Hypertext and Hypermedia	3
2009	Understanding the user: Personomy translation for <i>tag</i> recommendation	Wetzker R., Said A., Zimmermann C.	CEUR Workshop Proceedings	3
2009	Tags4Tags: Using tagging to consolidate <i>tag</i>	Garcia-Castro L.J., Hepp M., Garcia A.	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	3
2007	Mutual contextualization in tripartite graphs of folksonomies	Yeung C.-M.A., Gibbins N., Shadbolt N.	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	3
2013	Picture <i>tag</i> and world knowledge: Learning <i>tag</i> relations from visual semantic sources	Xie L., He X.	MM 2013 - Proceedings of the 2013 ACM Multimedia Conference	2

ANO	TÍTULO	AUTORES	TÍTULO DO JORNAL	SCOPUS
2012	Mining query log graphs towards a query folksonomy	Francisco A.P., Baeza-Yates R., Oliveira A.L.	Concurrency Computation Practice and Experience	2
2012	Detecting overlapping communities in folksonomies	Chakraborty A., Ghosh S., Ganguly N.	HT'12 - Proceedings of 23rd ACM Conference on Hypertext and Social Media	2
2012	Representing and visualizing folksonomies as graphs - A reference model	Lohmann S., Diaz P.	Proceedings of the Workshop on Advanced Visual Interfaces AVI	2
2012	Bridging informal tagging and formal semantics via hybrid navigation	Lezcano L., Garcia-Barriocanal E., Sicilia M.-A.	Journal of Information Science	2
2011	Exploring social relations for personalized <i>tag</i> recommendation in social tagging systems	Liu K., Fang B., Zhang W.	IEICE Transactions on Information and Systems	2
2010	HyperTwitter: Collaborative knowledge engineering via Twitter messages	Hepp M.	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	2
2010	Modeling ontology of folksonomy with latent semantics of <i>tag</i>	Daud A., Li J., Zhou L., Zhang L., Ding Y., Muhammad F.	Proceedings - 2010 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2010	2
2010	FolkDiffusion: A graph-based <i>tag</i> suggestion method for folksonomies	Liu Z., Shi C., Sun M.	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	2

ANO	TÍTULO	AUTORES	TÍTULO DO JORNAL	SCOPUS
2010	Incorporating relevance and importance for dynamic ranking in folksonomies	Liu K., Fang B., Zhang W.	Journal of Convergence Information Technology	2
2009	Relational classification for personalized <i>tag</i> recommendation	Marinho L.B., Preisach C., Schmidt-Thieme L.	CEUR Workshop Proceedings	2
2008	Locally expandable allocation of folksonomy <i>tag</i> in a directed acyclic graph	Eda T., Yoshikawa M., Yamamuro M.	Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	2

Fonte: do autor.

APÊNDICE D – TRABALHOS RELACIONADOS AO EGC

Nesse tópico é apresentada a bibliometria dos documentos do EGC relacionados a dissertação. Pesquisa realizada em janeiro de 2015 no Bando de Dados de Teses e Dissertações do Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento (PPGEGC).

Os quadros abaixo apresentam as palavras chaves de busca, o total de trabalhos encontrados e os mais adequados em relação à presente pesquisa, juntamente com o tema principal.

A pesquisa sobre “ideias” retornou 10 documentos. Os principais temas de pesquisado Programa estão relacionados à geração de Ideias por grupos criativos ou desenvolvimento da criatividade. No Quadro 16 são apresentadas as principais pesquisas por Prada (2009), Miguez (2012) e Dorow (2013).

Quadro 16 – Pesquisa sobre "ideias" no BTB do PPGEGC.

IDEIAS / IDÉIAS (10 documentos)	TEMAS
PRADA, Charles A. Proposta de modelo para o gerenciamento de portfólio de inovação: modelagem do conhecimento na geração de ideias. Dissertação, 2009.	Gestão de portfólio / geração de ideias
MIGUEZ, Viviane Brandão. Uma Abordagem de Geração de Ideias para o Processo de Inovação. Dissertação, 2012.	Geração de ideias
DOROW, Patrícia Fernanda. O Processo de Geração de Ideias para Inovação: Estudo de Caso em uma Empresa Náutica. Dissertação, 2013.	Geração de ideias

Fonte: do autor.

O Quadro 17 destaca os documentos de Ribeiro Junior (2011), Heinzle (2011) e Brignoli (2013). A pesquisa completa apresentou 29 trabalhos relacionados à “tomada de decisão” nas bases do PPGEGC. Esse tema se aproxima da necessidade de verificar os modelos de avaliação e seleção, então, assim se justifica o conhecimento em relação a presente pesquisa.

Quadro 17 – Pesquisa sobre "tomada de decisão" no BTB do PPGEGC.

TOMADA DE DECISÃO (29 documentos)	TEMAS
RIBEIRO JÚNIOR, Divino Ignácio. Modelo de sistema baseado em conhecimento para apoiar processos de tomada de decisão em ciência e tecnologia. Tese, 2011.	Tomada de decisão e Gestão do conhecimento
HEINZLE, Roberto. Um Modelo de Engenharia do Conhecimento para Sistemas de Apoio a Decisão com Recursos para Raciocínio Abduativo. Tese, 2011.	Cita Rede Semântica, mas não utiliza
BRIGNOLI, Juliano Tonizetti. Um Modelo para Suporte ao Raciocínio Diagnóstico diante da Dinâmica do Conhecimento sobre Incertezas. Tese, 2013.	Utiliza Redes Bayes

Fonte: do autor.

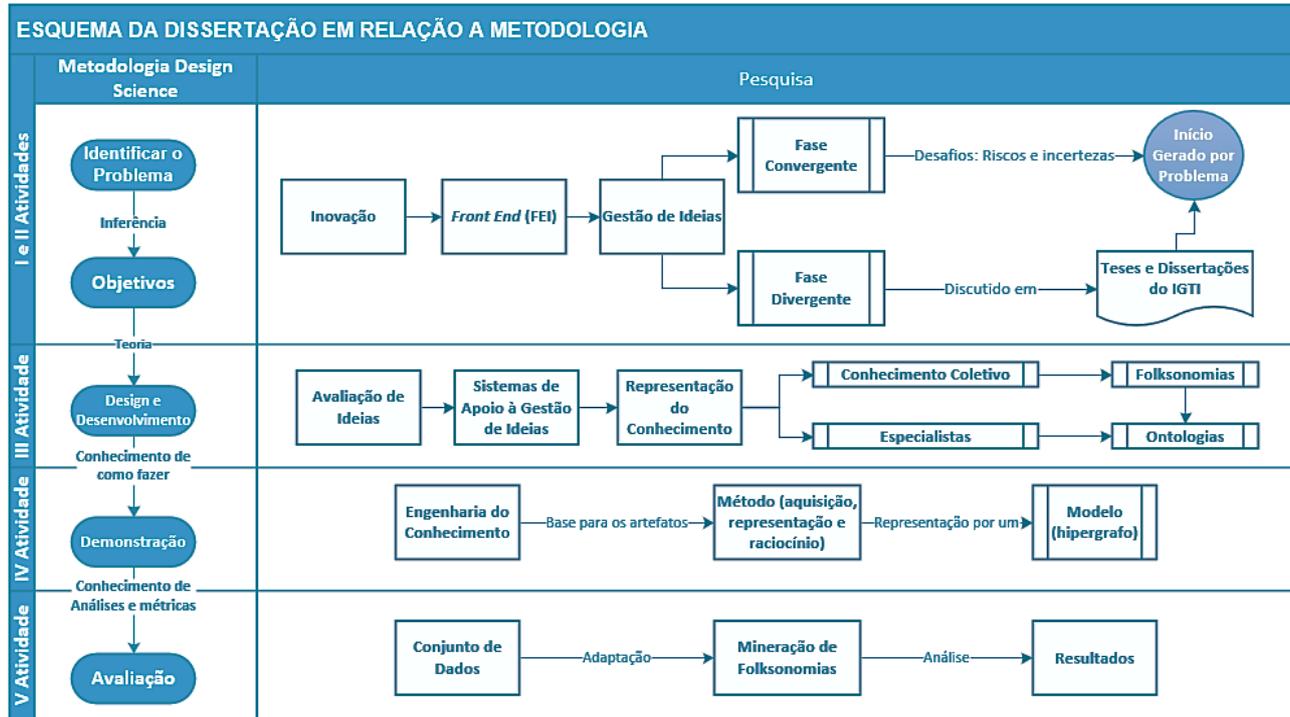
Os 19 documentos sobre “análise de rede” no banco de teses e dissertações do PPGEGC se referem mais as redes sociais. Porém, alguns documentos apresentam análises e discussões mais aprofundadas. É o caso de Balancieri (2010) e Fernandes (2012), como apresentado no Quadro 17

Quadro 18 – Pesquisa sobre "análise de rede" no BTB do PPGEGC.

ANÁLISE DE REDE (19 documentos)	TEMAS
BALANCIERI, Renato. Um método baseado em ontologias para explicitação de conhecimento derivado da análise de redes sociais de um domínio de aplicação. Tese, 2010.	Redes sociais
FERNANDES, Roberto Fabiano. Uma proposta de modelo de aquisição do conhecimento para identificação de oportunidades de negócio nas redes sociais. Dissertação, 2012.	Redes sociais

Fonte: do autor.

APÊNDICE E – ESQUEMA METODOLOGIA



Fonte: do autor.