



CENTRO UNIVERSITÁRIO LUTERANO DE PALMAS

ULBRA

COMUNIDADE EVANGÉLICA LUTERANA 'SÃO PAULO'
Recredenciado pela Portaria Ministerial nº 3.607 - D.O.U. nº 202 de 20/10/2005

Joabe Medeiros Gonçalves

**Definição de uma Ontologia para a Folksonomia de uma Rede Social
Acadêmica**

Palmas - TO

2011

Joabe Medeiros Gonçalves

**Definição de uma Ontologia para a Folksonomia de uma Rede Social
Acadêmica**

“Trabalho apresentado como requisito parcial da matéria Trabalho de Conclusão de Curso do Curso Sistemas de Informação, supervisionado pela Professora MSc. Parcilene Fernandes de Brito.”

Palmas - TO

2011

Joabe Medeiros Gonçalves

**Definição de uma Ontologia para a Folksonomia de uma Rede
Social Acadêmica**

“Trabalho apresentado como requisito parcial da matéria Trabalho de Conclusão de Curso do Curso Sistemas de Informação, supervisionado pela Professora MSc. Parcilene Fernandes de Brito.”

BANCA EXAMINADORA

Prof^a. MSc. Parcilene Fernandes de Brito

Centro Universitário Luterano de Palmas

Prof. MSc. Fabiano Fagundes

Centro Universitário Luterano de Palmas

Prof. MSc. Edeílson Milhomem da Silva

Centro Universitário Luterano de Palmas

Palmas - TO

2011

Agradecimentos

Agradeço primeiramente a Deus que está sempre dando oportunidades a mim e a minha família.

Agradeço em especial aos meus pais (Orlando e Deuza) pela preocupação e pelo amor que sempre tiveram por mim. Agradeço a eles também por sempre me apoiarem nas minhas decisões, por terem me dado e força e o suporte necessário, pelo carinho e orientação, enfim, por tudo que fizeram por mim. Essa caminhada não seria possível se não fosse por eles. Muito obrigado, Pai e Mãe.

Agradeço a minha irmã Midian que, apesar da distância, consegue estar presente sempre em minha vida, com vários minutos de conversa pelo celular e com muitas conversas sobre o cotidiano e sobre assuntos de “famílias estranhas” (Marte, E.T's etc).

Agradeço também ao meu irmão (primo) Laerth, que sempre esteve disponível para conversarmos sobre os nossos TCC's, e sobre histórias divertidas do nosso cotidiano. Obrigado meu irmão pelos incontáveis momentos de descontração, mesmo estando distante.

Agradeço a minha namorada Dulcimary, que esteve ao meu lado, sempre sabendo usar palavras para me confortar e incentivar em vários momentos. Agradeço sempre a Deus por ter me dado a oportunidade de tê-la conhecido nos corredores da Ulbra. Obrigado também por diversas vezes me ajudar a escolher os caminhos. Agradeço de coração.

Agradeço aos meus amigos do peito que tive oportunidade de conhecer na caminhada deste curso e que tenho certeza que essa amizade continuará ao longo da vida: Ricardo, Flávio, Heitor, Wesley, Warley (Macarrão), Warley (he-Man), Rafael (Rafiuski) e Wisley. Agradeço a vocês por ter transformado esses anos não só em aprendizado (com os estudos e vários trabalhos juntos), mas também em diversão. Agradeço também aos meus grandes e primeiros amigos de Palmas, Pablo (Peixinho), Diogo (“tchieu” irmão), Danilo (Japa “oopaa”), Fernando e Samuel, que são amigos que sempre tem motivos para descontrações, insistências e o mais importante, companheirismo. Obrigado meus amigos.

Agradeço a minha orientadora Parcilene pela paciência que teve comigo desde o início do TCC, e pelas palavras que muitas vezes eram dramáticas em relação às correções. Eu muitas vezes acabava me divertindo com os seus comentários nas correções. Agradeço pelo conhecimento compartilhado e pelas incontáveis explicações de como seguir. Obrigado Parcilene.

Agradeço ao meu professor Fabiano, que sempre me incentivou a seguir adiante no curso, sempre teve tempo para conversar sobre trabalhos, mestrados, empresas, enfim, sempre compartilhou o seu conhecimento. Agradeço também por ter me indicado para o meu primeiro emprego na área. A partir desse trabalho muitas portas se abriram. Agradeço pela confiança ao me indicar. E agradeço a todos os professores: Jackson, Mádía, Fernando, Andrés, Cristina, Edeílson e Michael.

Enfim, agradeço a todos os meus familiares e amigos que acreditaram em mim. Obrigado a todos.

RESUMO

Atualmente há um número crescente de redes sociais que utilizam os conceitos do processo de marcação através das *tags* e os conceitos da folksonomia. Mas esses sistemas baseados em folksonomia têm alguns problemas que podem confundir os usuários, como palavras ambíguas, sinônimos, falta de interoperabilidade entre os sistemas etc. Com uma ontologia de *tags* aplicada em uma rede social acadêmica, há um compartilhamento de informações entre os usuários. Através da análise das nuvens de *tags* contextuais, pode ser verificada a similaridade entre duas ou mais nuvens de *tags* e, assim, pode haver sugestões de conteúdos e compartilhamento da informação entre os usuários da rede social. Nesse trabalho serão apresentados os conceitos de uma ontologia de *tags* e a aplicação dessa ontologia em uma rede social acadêmica.

Palavras-chave: folksonomia, *tags*, ontologia, *tagging ontology*.

Lista de Figuras

Figura 1. Conexões em uma rede social. (<i>Fonte:</i> http://static.howstuffworks.com/gif/myspace-networking.jpg).....	10
Figura 2. Grafos rodoviário dos Estados Unidos.(RECUERO, 2006, p. 28)	11
Figura 3. 12 Princípios das comunidades. (SILVA, 2005).....	13
Figura 4. Organização através de especialistas. (PEREIRA, 2008)	15
Figura 5. Organização com Folksonomia. (WAL, 2006)	17
Figura 6. Retorno dos dados em uma busca.	18
Figura 7. Personomia de um usuário. (Silva, 2009)	18
Figura 8. Conjunto de personomias forma uma folksonomia	19
Figura 9. Processo de Tagging.	20
Figura 10. Falta de interoperabilidade das folksonomia atualmente. (<i>Fonte:</i> Knerr, 2006).....	28
Figura 11. Interoperabilidade dos sistemas. (<i>Fonte:</i> Knerr, 2006)	29
Figura 12. Relacionamentos. (<i>Fonte:</i> DAMME, 2007).....	30
Figura 13. Ambiguidades nos sistemas atuais. (<i>Fonte:</i> PASSANT e LAUBLET, 2008)	31
Figura 14. Ontologia de Tags de Knerr. (<i>Fonte:</i> KNERR, 2006).....	32
Figura 15. Representação proposta por Gruber.	34
Figura 16. Junção de ontologia com Processo de Tagging. (<i>Fonte:</i> GONÇALVES, 2006)	35
Figura 17. Folknomia + Ontologia de <i>Tags</i>	41
Figura 18. Estrutura do Konnen	42
Figura 19. Representação do esquema de marcação na visão do usuário e criador do conteúdo.	42
Figura 20. Visualização da Nuvem de <i>tags</i>	43
Figura 21. Tempo X Marcações de Objetos.	44

Figura 22: Sugestão por Similaridade de Tags.....	45
Figura 23. Elementos do processo de <i>tagging</i> para o Konnen	46
Figura 24. Estrutura do banco de dados do processo de <i>tagging</i> do <i>Konnen</i>	47
Figura 25. Estrutura das <i>Views</i>	50
Figura 26. Estrutura <i>ViewAnnotationsByAnoSemestre</i>	51
Figura 27. Nuvem de <i>Tags</i> simples por comunidade	52
Figura 28. Nuvem de <i>Tags</i> por comunidade baseado na polaridade	53
Figura 29. Processo de <i>Tagging</i>	53
Figura 30. Classe <i>TagResultsItem</i>	54
Figura 31. Construtor <i>TagResultItem</i>	56
Figura 32. Método <i>ResultViewTagCloud</i>	56
Figura 33. Método <i>ResultViewAnnotation</i>	56
Figura 34. Construtor para a <i>ViewUserAnnotation</i>	57
Figura 35. Método <i>TagCloudByCommunityId</i>	57
Figura 36. Método <i>TagCloudPolarityByContentId</i>	58
Figura 37. Gráfico da Similaridade (Fonte: Wives, 2002, p. 39)	59
Figura 38. Cálculo do ângulo de similaridade. Fonte: (Wives, 2002, p. 40).....	59
Figura 39. Método <i>OrganizeList</i>	60
Figura 40. Método <i>CommunitySimilarity</i>	61
Figura 41. Nuvem de <i>tags</i> por comunidade.....	62
Figura 42. Nuvem de tags por comunidade baseada na polaridade.....	63
Figura 43. Função SQL <i>fnViewResultsPolarityCommunity</i>	64
Figura 44. Retorno XML da função <i>fnViewResultsPolarityCommunity</i>	64
Figura 45. Ontologia de Tags para o Konnen.....	65

Sumário

1	INTRODUÇÃO	7
1.1	Organização do Trabalho.....	8
2	REVISÃO DE LITERATURA	10
2.1	Redes Sociais.....	10
2.2	Taxonomia em Conteúdos da Web.....	14
2.3	Folksonomia.....	16
2.3.1	Processo de Tagging e Componentes.....	20
2.3.2	Problemas com Folksonomia e Semântica da Web	21
2.3.3	Considerações.....	22
2.4	Ontologia.....	22
2.4.1	Características.....	24
2.4.2	Objetivos	24
2.4.2.1	Ontologias Simples	25
2.4.2.2	Ontologias Estruturadas	25
2.4.3	Componentes de uma ontologia.....	26
2.5	Ontologia de Tags (<i>Tagging Ontology</i>).....	27
2.5.1	Características e Objetivos.....	30
2.5.2	Representação e Elementos Principais.....	32
2.5.3	Aplicação em uma Rede Social	34
3	MATERIAL E MÉTODOS	37
3.1	Período.....	37
3.2	Materiais	37
3.3	Softwares.....	37
3.4	Metodologia.....	38
4	RESULTADOS E DISCUSSÃO	40
4.1	Contextualização do domínio	40
4.2	Estrutura do Konnen e Visualização das Nuvens	41
4.3	Processo de Tagging no Konnen	46
4.4	Estrutura do Banco de Dados	47
4.5	Requisitos	48
4.6	Aplicação	49

4.6.1 Views	49
4.6.2 Implementação	54
4.6.3 Nuvens de Tags.....	62
4.6.4 Retorno XML	63
4.7 Conclusão	65
5 CONSIDERAÇÕES FINAIS	67
6 REFERÊNCIAS	69

1 INTRODUÇÃO

As redes sociais na internet se apresentam como um grande atrativo devido às conexões feitas entre as pessoas com o intuito de fazer novas amizades, participarem de comunidades, divulgação de seus trabalhos e/ou busca de novos conhecimentos. As redes sociais, como cita Recuero (2006, p. 14), “são uma rede de atores formada pela interação social mediada pelo computador”. A maioria das redes sociais é focada em laços de amizades e de popularidade, tais redes proporcionam uma visão detalhada dos relacionamentos.

Já em algumas redes sociais em que o foco é a distribuição do conhecimento, como fóruns e *wikis*, alguns usuários utilizam marcações ou *tags* para compartilhar o conteúdo marcado ou para uso próprio. As *tags* são marcações, atribuições de valores aos conteúdos disponíveis na web. Nesses ambientes de redes sociais, a participação dos usuários é tão ativa que eles mesmos fazem a organização e hierarquização de seus conteúdos através dessas *tags*, participando de formulação de novas regras e novos conteúdos, como é o exemplo de fóruns de comunidades. Segundo Costa *et al* (2008, p. 196) “comunidades podem ser vistas como agregações de pessoas com objetivos em comum”, em que os usuários buscam conhecimentos baseados em seus interesses.

Em meio a uma troca constante de conhecimentos e discussões feitas nessas comunidades, o usuário pode se perder em qual discussão participou, qual tópico achou relevante, e/ou até mesmo ter dificuldade em encontrar pessoas com objetivos e interesses em comum. Para tentar sanar parte desses problemas, algumas redes sociais deram a liberdade para os usuários marcarem seus conteúdos com as *tags*.

Nos sistemas baseados em folksonomia, o usuário tem a liberdade de fazer sua própria organização e distribuir os conteúdos através das marcações. A palavra folksonomia é um termo criado por Thomas Vander Wal que é a junção das palavras “*folks*” que significa pessoas e “*taxonomia*”, que quer dizer a marcação, hierarquização dos conteúdos. Aquino (2008, p. 307) discorre que “a folksonomia ocorre através da atividade de seus usuários, que ficam livres para representar e recuperar os dados com base no senso comum”, com isso a folksonomia torna-se uma forma eficiente de organização dos conteúdos na web, e essas organizações muitas vezes fazem surgir novos métodos de organização para outros usuários.

Em sistemas baseados em folksonomia não há um modo padrão sobre como organizar os conteúdos, fazer sugestões de outras marcações, e/ou fazer ligações entre marcações e objetos de uma maneira automática. Assim, para que seja possível sanar alguns problemas,

pode-se definir uma ontologia baseada nas marcações feitas pelos usuários nesses sistemas. Vilar (2006, p. 35) define ontologia como uma conceitualização de parte de um domínio e o compartilhamento das informações deste domínio para que o entendimento possa ser analisado com facilidade pelas pessoas e processado automaticamente pelo computador. Através da aplicação de uma ontologia em uma rede social, os usuários poderão ter um *feedback* mais eficiente no que tange a distribuição e compartilhamento dos objetos presentes nessa rede.

Ao fazer a junção dos conceitos de folksonomia e os conceitos de ontologia forma-se um conceito chamado *tagging ontology* ou ontologia de *tags*. Tal conceito baseia-se no aprendizado e inferências feitas dentro de um domínio, nas marcações que os usuários fazem nos objetos.

O objetivo deste trabalho é apresentar o estudo bibliográfico necessário para a definição das etapas relacionadas ao desenvolvimento de uma ontologia baseada na folksonomia de uma rede social. Como também apresentar a ontologia de *tags* aplicada na rede social acadêmica. Neste trabalho serão apresentadas as etapas para a correta execução do desenvolvimento dessa ontologia de *tags*, apresentando desde os requisitos do sistema até a implementação final. A partir disso, o sistema poderá explicitar o significado das *tags* aos usuários, apresentar diversos tipos de nuvens de *tags* e verificar a similaridade das comunidades.

1.1 Organização do Trabalho

O trabalho está organizado da seguinte forma:

- Capítulo 2 (Revisão de Literatura) – Na seção 2.1 é apresentada uma breve introdução às redes sociais. A seguir (seção 2.2) é apresentada a definição de taxonomia e como era usado tal conceito na web. Na seção 2.3 são apresentados os conceitos de folksonomia, alguns sistemas que utilizam tais conceitos e uma discussão sobre alguns problemas que envolvem folksonomia. Na seção 2.4 são discutidos os conceitos de ontologia, subdividindo tal seção em características, objetivos, componentes, tipos de ontologia e estudo de criação de uma ontologia. A seguir são apresentados e discutidos os conceitos de ontologia de *tags*, a seção está subdividida em características, objetivos, representação e elementos, e aplicação em uma rede social. Tal seção subdivide-se em: características e objetivos, representação e elementos principais, e aplicação em uma rede social.

- Capítulo 3 (Materiais e Métodos) – Apresenta os materiais e a metodologia utilizada para o desenvolvimento do trabalho.
- Capítulo 4 (Resultados e Discussão) – Na seção 4.1 é apresentada uma breve contextualização do Konnen. A seguir (seção 4.2) é explicado sobre a estrutura do Konnen e como são definidas as visualizações dados nas nuvens de *tags*. Na seção 4.3 é estudado e mostrado o processo de marcação dentro do *konnen*. Na seção 4.4 é apresentada a estrutura do banco de dados. Já na seção 4.5 são discutidos os requisitos que englobarão o sistema. A seção 4.6 discorre sobre a aplicação em si, e tal seção subdivide-se em quatro outras seções:
 - *Views* (seção 4.6.1) – nessa seção são apresentadas as diversas *views* que retornam dados das nuvens de *tags*.
 - Implementação (seção 4.6.2) – é a seção que discute a parte da implementação. Inicialmente (seção 4.6.2.1) é apresentada a classe que estrutura o objeto baseado na consulta feita no banco de dados. A seção 4.6.2.2 discorre sobre a classe que busca as diversas nuvens de *tags* no banco. Já a seção 4.6.2.3 apresenta os métodos necessários para fazer o cálculo de similaridade entre duas comunidades.
 - Nuvens de *Tags* (seção 4.6.3) – é mostrada uma exemplificação visual de como são as nuvens de *tags*.
 - Retorno XML (seção 4.6.4) – é apresentado um modelo de função que retorna os dados de uma nuvem de *tags* em um documento XML estruturado.
- Capítulo 5 (Considerações Finais) – São apresentadas as conclusões obtidas com o desenvolvimento deste trabalho através dos estudos realizados.
- Capítulo 6 (Referências) – São apresentadas as referências utilizadas no decorrer do trabalho.

2 REVISÃO DE LITERATURA

Nesta seção serão apresentados os conceitos que envolvem as redes sociais, taxonomias, *tags*, sistemas baseados em folksonomia e ontologia. Além disso, serão discutidos os elementos que advém da junção dos conceitos de ontologia e folksonomia.

2.1 Redes Sociais

Na internet, o termo “Redes Sociais” tornou-se parte do cotidiano dos usuários, e segundo Belleghem (2010, p. 8) 90% desses usuários tem o conhecimento de pelo menos uma rede social. Tal termo está popular, tanto no mundo corporativo (com as redes sociais internas, ou seja, redes sociais feitas e usadas somente dentro das corporações para uso dos funcionários) como para os usuários finais (como as redes sociais de amizades, contatos profissionais, comunidades etc). As redes sociais, segundo Recuero (2006, p.14), “são as redes de atores formadas pela interação social mediada pelo computador”. Na Figura 1 é apresentada a representação das conexões entre os atores de uma rede social.

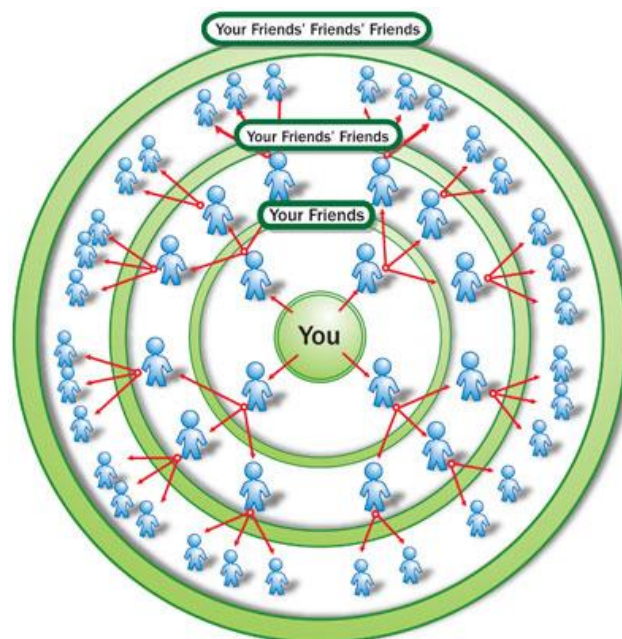


Figura 1. Conexões em uma rede social. (Fonte: <http://static.howstuffworks.com/gif/myspace-networking.jpg>)

Os sites de redes sociais focados em laços de amizades e popularidade tornam “visíveis as redes invisíveis”, que são as conexões e laços de amizades que podem existir entre os usuários, através de sugestões e análise de fotos, detalhamento de contatos e *links*.

De acordo com Roos (online) o vínculo social em uma rede pode ser dividido em dois, que são: os laços fortes e laços fracos, ou seja, nos laços fortes estão presentes a família e amigos íntimos, enquanto nos laços fracos estão presentes pessoas que se conheceram em situações específicas e não precisam manter um contato frequente. Recuero (2006, p. 57) cita que, uma vez que as conexões entre as pessoas em uma rede social são estabelecidas, elas sempre estarão presentes, independente do tempo.

A pesquisa sobre redes sociais perpassa alguns objetivos, a saber:

- estudar a formação das redes sociais humanas, ao estudar como os laços se formam, e qual o vínculo que as pessoas têm, por exemplo, laços de interesse comercial, amizade.
- Como ressalta Costa *et al* (2008, p. 192), tais redes servem também para identificar dificuldades de comunicação de um grupo no interior de uma organização através dos vínculos de amizade e de colaboração entre eles. Com as redes sociais nas organizações, o gerenciamento e compartilhamento do conhecimento pelos funcionários, tem se demonstrado bastante eficiente, quando aplicado corretamente.

As redes sociais se dividem em três tipos, que são as redes igualitárias, redes de mundos pequenos e redes sem escala. De acordo com Recuero (2006, p. 28) alguns especialistas em estudos de redes igualitárias dizem que nessas redes os nós possuem aproximadamente o mesmo número de conexões, ou seja, a quantidade de nós aumenta proporcionalmente igual à medida que outros nós da rede aumentam. Um exemplo destas redes igualitárias é a rede rodoviária dos Estados Unidos apresentada na Figura 2.

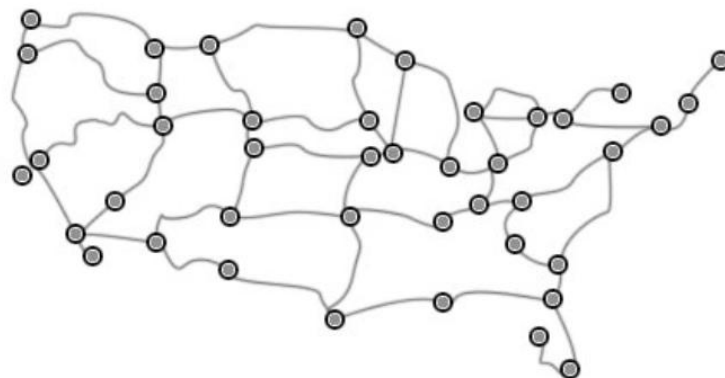


Figura 2. Grafos rodoviário dos Estados Unidos.(RECUERO, 2006, p. 28)

De acordo com o apresentado na Figura 2, a rede rodoviária dos Estados Unidos apresenta aproximadamente a mesma quantidade de nós, para que se possa facilitar a integração e as partidas com as cidades próximas, formando assim uma rede de conexões igualitárias.

Já em redes de mundos pequenos as pessoas estão a poucos graus de separação umas das outras, que pode ser entendido pela quantidade de nós que separam os usuários que não tenham uma conexão direta, ou seja, usuários que não “se conhecem”. Esse modelo é o que está sendo utilizado nas redes sociais atuais, em que, quando um nó (usuário) se conecta com outro, a distância ou o grau de separação com outros usuários diminui (RECUERO, 2006, p. 30-33). Em redes sociais como o *LinkedIn*, *Facebook*, *Orkut* etc, tem-se a sugestão de amigos, e tais sugestões funcionam a partir da conexão que uma pessoa tem com outras pessoas indiretamente, ou seja, baseado no grau de separação entre as pessoas.

As redes sem escalas podem ser compreendidas como as redes de citações ou de publicações, que segundo Recuero (2006, p. 34) um autor renomado pode ter 80% das conexões de uma comunidade. Nessas redes quanto mais conexões um nó possui, maiores chances de ele ter mais novas conexões como cita Recuero (2006, p. 34). Como é um exemplo de um usuário no *Twitter* que, muitas vezes, quanto mais seguidores ele tiver, maiores são as chances de acumular ainda mais seguidores.

Outro conceito de rede social são as comunidades virtuais, em que as pessoas se reúnem para discutir sobre os mais variados assuntos e que segundo Costa *et al* (2008, p. 196), "as comunidades virtuais podem ser vistas como agregações de pessoas com objetivos em comum", ou seja, os usuários buscam as comunidades baseadas em seus interesses, como é o caso do *Orkut*, que as pessoas buscam por diversas comunidades de diversos temas. Tais comunidades são as principais fontes de distribuição do conteúdo entre os usuários. Algumas vantagens podem ser percebidas nessas comunidades virtuais, que são:

- possibilidade de conhecer novas culturas.
- trocar informações.
- aprender de forma coletiva.

Na Figura 3 são apresentados exemplos de alguns componentes das comunidades virtuais.



Figura 3. 12 Princípios das comunidades. (SILVA, 2005)

Os doze componentes apresentados na figura 3, que englobam as comunidades virtuais, são (SILVA, 2005):

- História que a comunidade tem, se é antiga ou não.
- A liberdade de expressão que as comunidades devem disponibilizar.
- A troca de informações entre os integrantes.
- A confiança das informações passadas.
- Não pode haver fronteiras, pois são comunidades virtuais e podem ser acessadas de qualquer lugar.
- Um ambiente que estimule os usuários a participar ativamente.
- Os grupos formados dentro das comunidades para discussão de assuntos ainda mais específicos.
- A comunicação entre os usuários da comunidade deve ser facilitada.
- De livre acesso, sem a intervenção do governo.
- Depende sempre da reputação que tem.
- O foco da comunidade, ou seja, como as pessoas acharão a comunidade e porque irão procurar.
- E por último, o objetivo da comunidade, ou seja, qual a finalidade da comunidade, para quais discussões e debates é útil.

Tais componentes são analisados para verificar se a comunidade é ativa ou não, se as informações são de livre acesso, se há liberdade de expressão, foco e objetivo bem definidos. Quanto mais desses componentes houver em uma comunidade, maior será a contribuição dos usuários na gestão e distribuição do conhecimento, pois há uma facilitação na participação dos usuários.

As redes sociais e suas comunidades são ferramentas úteis no campo de gestão do conhecimento e úteis também em sistemas de recomendação (CRUZ, p. 20). No que tange a gestão do conhecimento dentro de ambientes organizacionais, as redes sociais funcionam como meio de distribuição e compartilhamento do conteúdo entre diversos setores da empresa (TOMAÉL et al, 2005, p. 94), fazendo surgir o compartilhamento do conhecimento através das comunidades, foruns etc. Por exemplo, quando um usuário faz marcações de fácil entendimento para o setor no qual trabalha podem surgir novas ideias, métodos de trabalho etc.

Nas redes sociais as informações transitam em um fluxo contínuo, já que essas redes incentivam o compartilhamento de informações através de comunidades, marcações e fóruns. Isso faz com que os usuários participem na formulação, organização e divulgação dos conteúdos, incentivando assim a colaboração e o aprendizado, ou seja, os usuários utilizam os conceitos de inteligência coletiva. Segundo Johnson (2001, p. 7-8) a inteligência coletiva é “a capacidade de uma organização, comunidade ou grupo, de fazer perguntas e buscar as respostas juntos”. Percebe-se que tal conceito está presente, em um dado nível, nos *blogs*, *wikis*, projetos *open source*, comunidades virtuais etc.

2.2 Taxonomia em Conteúdos da Web

Uma taxonomia é a hierarquização/organização das informações feitas seguindo um modelo pré-definido e controlado por pessoas especializadas, como, por exemplo, nos sistemas mais antigos que utilizavam marcações exclusivas e hierárquicas para a organização de seu conteúdo, sua estruturação era exclusiva e inflexível, pois o conteúdo e as marcações eram organizados por especialistas e quando o conteúdo era exposto aos usuários toda organização já estava estabelecida (GOLDER E HUBERMAN, 2006, p. 198). Esses sistemas com as categorias pré-definidas permitiam que os usuários realizassem uma busca ou uma inserção de um conteúdo através das categorias existentes, ou seja, o usuário não podia buscar ou inserir suas marcações de acordo com o seu entendimento. Por exemplo, “laranja” e “limão” são classificados como frutas cítricas no sistema pelos especialistas, sendo que tais frutas fazem

parte da mesma região e um usuário gostaria de marcar também como “frutas tropicais” para melhor entendimento futuro, no entanto, em um sistema baseado somente em taxonomia, esse usuário não podia fazer sua marcação ou, se pudesse, ao realizar a marcação das frutas tropicais o sistema não deixaria que ele marcasse também como frutas cítricas. Desta forma, as hierarquizações impossibilitavam uma maior flexibilidade no sistema, pois a taxonomia organiza ou categoriza os conteúdos através de abordagens formais controladas por especialistas.

Uma representação sobre como são criados e categorizados os conteúdos em um sistema baseado em taxonomia é apresentado na Figura 4 (PEREIRA, 2008, p. 40).

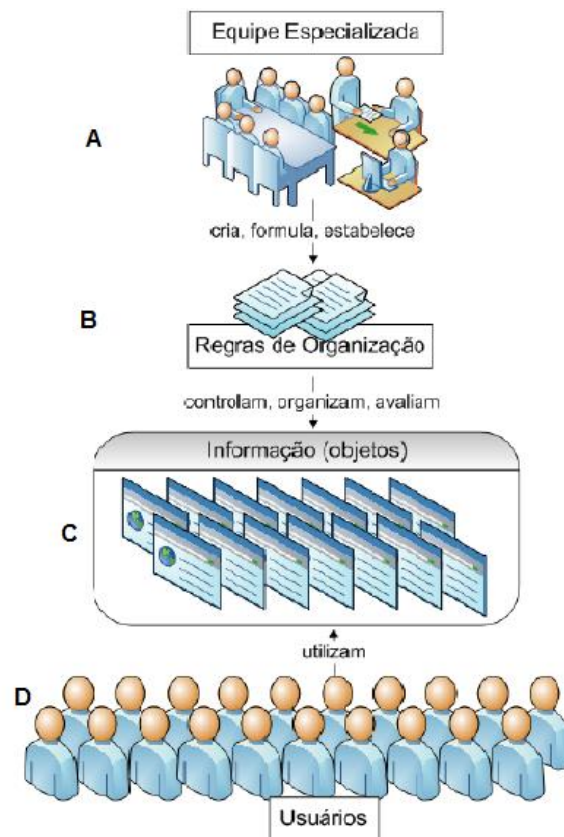


Figura 4. Organização através de especialistas. (PEREIRA, 2008)

Na Figura 4 observa-se a formulação de conteúdos na web através de especialistas, seguindo etapas como:

- **A:** A equipe de profissionais especializados se reúne para criação do conteúdo;
- **B:** Estabelecem regras, criam marcações pré-definidas e as lançam;
- **C:** Através das regras de organização, os especialistas controlam o conteúdo, avaliam e assim disponibilizam para os usuários;

- **D:** Os usuários podem somente ver os conteúdos e marcá-los com um vocabulário pré-definido.

Segundo Golder e Huberman (2006, p. 198), em sistemas com taxonomia exclusiva e hierárquica, a busca é genérica, ou seja, o usuário informa o nome exato do que procura, mas faz uma busca do conteúdo procurado navegando entre as hierarquias pré-definidas, o que (dependendo da organização, quantidade de conteúdo e nomes culturais) pode levar tempo para achar algum conteúdo relevante. Com o crescimento no número de usuários nos sistemas que utilizavam taxonomia e a grande quantidade de conteúdos lançada diariamente, surgiu uma necessidade de uma organização diferenciada e assim o modelo em que os especialistas faziam o controle dos conteúdos passou a ser inviável, pois, segundo Mathes (2004, online), esse controle pode ser de alta qualidade, mas requer muito tempo e esforço, além de ser pouco flexível e com mínima interação com os usuários. Surgiram assim sites em que o usuário pode fazer suas próprias marcações e organizar o conteúdo da forma que for mais apropriada ao seu contexto. Esses sistemas são chamados sistemas baseados em folksonomia, que serão apresentados na próxima seção.

2.3 Folksonomia

O termo folksonomia foi criado por Wal (2005) com intuito de explicar que folksonomia é uma taxonomia definida por qualquer usuário de um sistema. Desta forma, as marcações e formulações dos conteúdos são feitos de forma colaborativa, em que os usuários utilizam uma linguagem comum de acordo com a comunidade que utilizará o conteúdo, então percebe-se que partir de tal conceito a própria comunidade cria os seus padrões e marcações.

O'Reilly (2005, online) cita que a folksonomia é a evolução da taxonomia, pois a organização, classificação e distribuição dos conteúdos lançados na web não são limitados a um grupo de especialistas. Já Wal (2005, online) define a folksonomia como “o resultado da marcação pessoal de qualquer conteúdo com URL para uso próprio”. Atualmente, alguns sites utilizam os conceitos da folksonomia para melhoria da organização de seus conteúdos, já que permite a cada usuário fazer sua própria organização, como é o exemplo do *Delicious*, *Flickr*, *Youtube* etc.

Na Figura 5 tem-se uma representação de distribuição do conteúdo e sua organização com o uso de folksonomia.

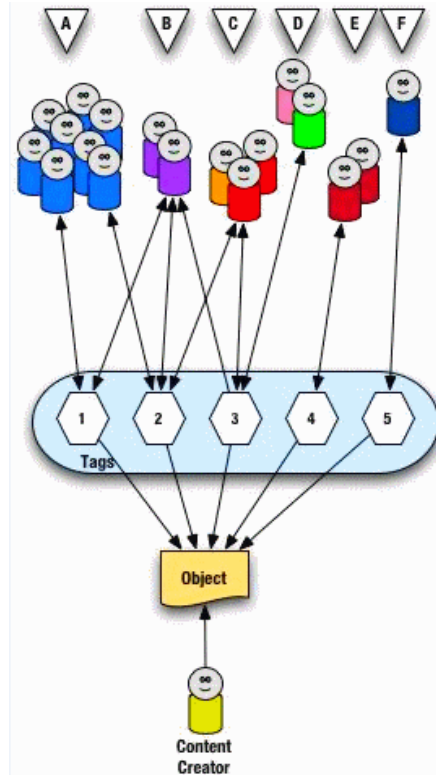


Figura 5. Organização com Folksonomia. (WAL, 2006)

De acordo com a Figura 5, os usuários fazem a criação (*content creator*), avaliação, categorização dos objetos lançados na web e atribuem sentido para estes objetos com as *tags*. Os objetos categorizados são utilizados por outros usuários, de outras comunidades e/ou sistemas, que são representados pelas letras A, B, C, D e F. Os usuários podem também utilizar as marcações de outros usuários como fonte de busca. De acordo com Pereira (2008, p. 40), após a atribuição de sentido e marcações feitas pelos usuários, pode surgir um padrão a ser seguido pela comunidade, formas de organização e compartilhamento do conteúdo.

Em um ambiente cujo contexto é embasado nos conceitos de folksonomia, os usuários ficam livres para usar, recuperar e referenciar os dados da maneira que achar melhor. Isso faz com que a web seja dinâmica, baseada no interesse dos usuários, sendo os conteúdos atualizado por eles, assim, a folksonomia tem um *feedback* imediato (AQUINO, 2008, p. 306). O *feedback* imediato está presente tanto nas buscas, em que o usuário tem uma busca mais específica, tanto nas próprias marcações, em que o usuário pode fazer modificações de suas marcações e buscar marcações semelhantes de uma forma livre. Em buscas de conteúdos, tem-se um retorno mais específico, pois ele pode marcar e também visualizar marcações personalizadas e, muitas vezes, mais relevantes, como se pode observar na Figura 6.

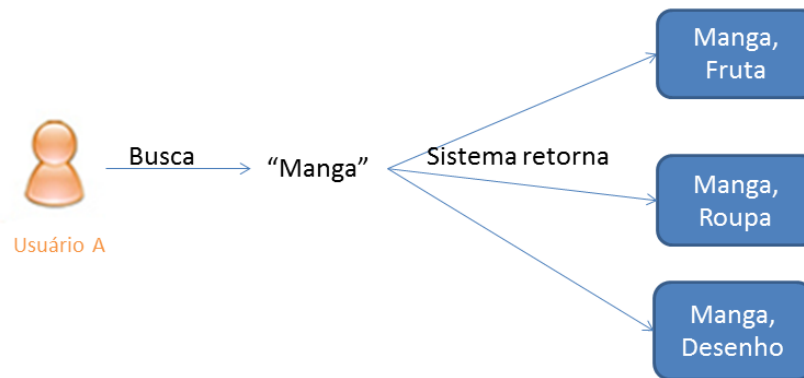


Figura 6. Retorno dos dados em uma busca.

Segundo a representação da Figura 6, um usuário ao procurar pela palavra “manga” no sistema, terá como retorno conteúdos que foram marcados com “manga” e também serão apresentadas as classificações que foram feitas juntamente com o nome buscado. Assim, o usuário poderá ver as marcações “manga, roupa”, “manga, fruta” ou “manga, desenho”, o que retira a ambiguidade da pesquisa e permite uma busca mais específica. Desta forma, as *tags*, além de melhorar as buscas pelos conteúdos, podem ajudar a criar uma aproximação com os conteúdos relacionados a partir das palavras-chaves.

Silva (2009, p. 37) cita que a folksonomia é um conjunto de várias personomias em uma rede social. Uma personomia pode-se entender que é um conjunto de *tags* e marcações feitas por um único usuário na rede.

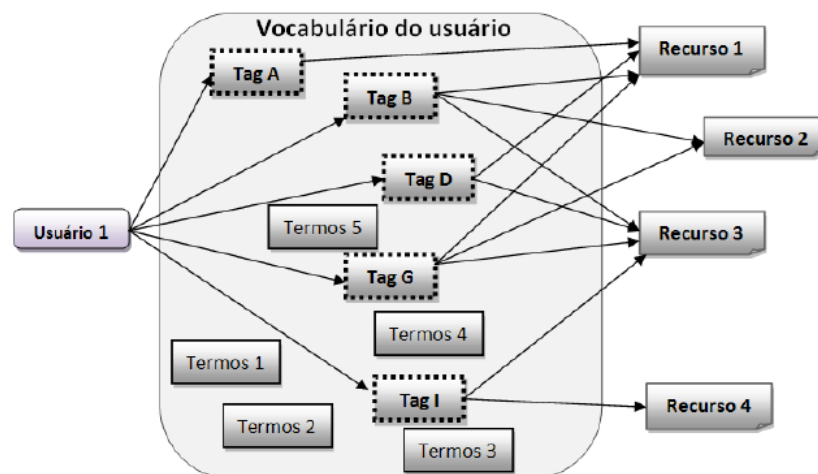


Figura 7. Personomia de um usuário. (Silva, 2009)

De acordo com a Figura 7, um usuário faz as marcações dos objetos (recursos) com um vocabulário comum para ele ou para sua comunidade a partir de termos que possam ser relevantes para o contexto, possibilitando assim uma busca eficiente desses objetos. Essas

marcações passam a ser representadas em sua nuvem de *tags*. Já na Figura 8, tem-se um conjunto de personomias.

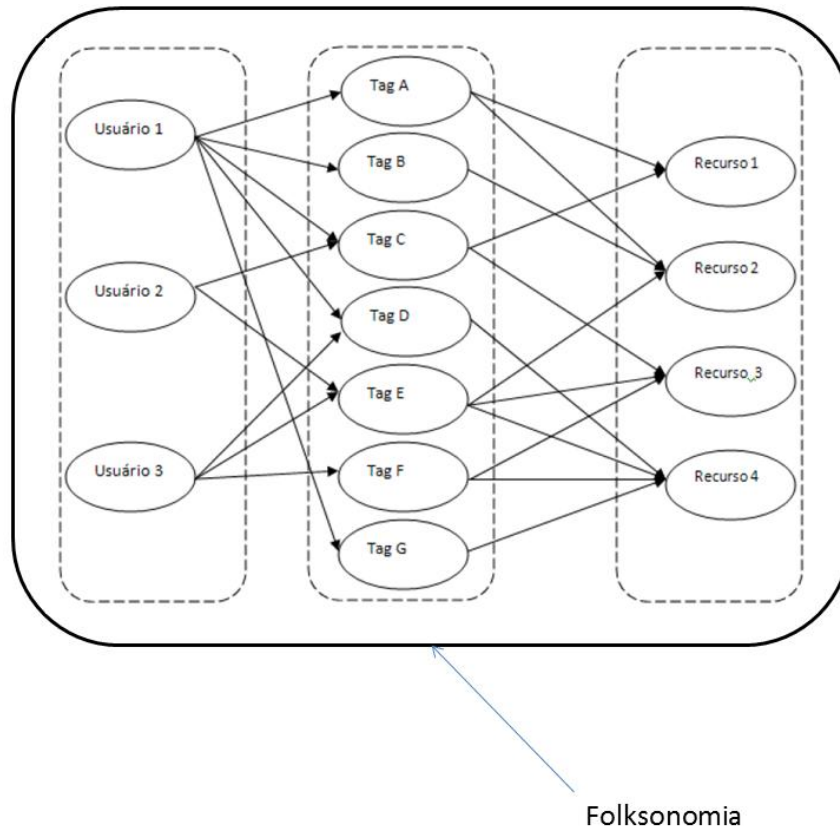


Figura 8. Conjunto de personomias forma uma folksonomia

A Figura 8 ilustra um conjunto de usuários realizando as marcações em diversos conteúdos em um sistema. De acordo com a ilustração, cada usuário tem sua personomia, formando assim a folksonomia daquele sistema.

Os sistemas baseados em folksonomia, além de disponibilizar liberdade aos usuários em sua organização, tem como principais objetivos (MATHES, 2004, online):

- Distribuição e expansão do conhecimento;
- Facilitação nas buscas por conteúdos;
- Facilitação na organização da estrutura da web;
- *Feedback* imediato aos usuários.

Atualmente esses conceitos que envolvem folksonomia estão sendo cada vez mais utilizados em redes sociais para organização, busca e distribuição do conteúdo. Com a marcação ativa dos usuários nos conteúdos disponíveis na web, esses sistemas utilizam processo de *tagging*, que representa todo o processo de marcação, categorização e divulgação de um conteúdo através dessas *tags*.

2.3.1 Processo de Tagging e Componentes

O processo de *tagging* ou processo de marcação de acordo com Kim et al (2008, p. 130) é um processo que permite e estimula os usuários a compartilhar qualquer conteúdo em redes sociais. Nos sistemas baseado em folksonomia, a marcação dos conteúdos se faz a partir desse processo, e segundo Silva (2009, p. 31) ele é efetuado de forma coletiva, sendo os usuários os responsáveis principais para a efetuação das marcações. Guy e Tonkin (2006, online) citam que em sistemas baseados em folksonomia há pessoas que contribuem para as marcações e há outros que são os consumidores das marcações, ou seja, os sistemas são divididos em dois grupos, os colaboradores e os consumidores.

O processo de marcação dos conteúdos tem como característica a liberdade para o usuário fazer suas categorizações, sem um limite de *tags* que ele possa usar. Segundo Wal (2005, online), nesse processo há três pontos principais, como apresenta a Figura 9.

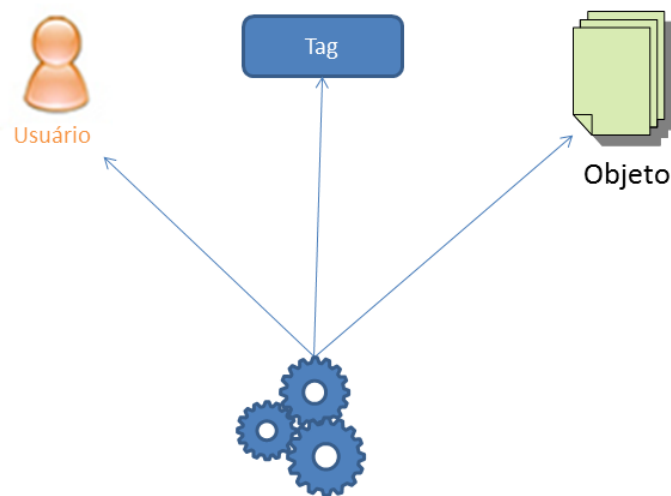


Figura 9. Processo de Tagging.

De acordo com a Figura 9, o processo de marcação é composto por:

- o usuário que é responsável por fazer a marcação;
- o conjunto de *tags* utilizado para referenciar um conteúdo;
- o objeto ou conteúdo que está sendo referenciado com o conjunto de *tags*.

Esses pontos são importantes para a recuperação dos conteúdos na web, pois um usuário poderá realizar sua busca a partir do conjunto de *tags* ou a partir das referências de um outro usuário. O processo de *tagging* é também um processo importante para estimular outros

usuários a organizar o sistema. Ao realizar as marcações com uma linguagem simples e fácil, o usuário estará contribuindo com a busca dos dados no sistema por outros usuários.

A partir da liberdade que o usuário tem para marcar os conteúdos da web nos sistemas baseados em folksonomia, podem surgir alguns problemas para os próprios usuários, que serão explicados na seção seguinte.

2.3.2 Problemas com Folksonomia e Semântica da Web

De acordo com Guy e Tonkin (2006, online) “nos sistemas baseados em folksonomia, qualquer usuário pode adicionar suas *tags* livremente, com isso pode ocorrer ambiguidade”, em muitos casos, usa-se também de marcações pessoais, tornando assim as buscas por novos conteúdos através das marcações ineficazes. Mathes (2004, online) cita os principais problemas da folksonomia, a saber:

- **Tags com significados pessoais:** As *tags* com significados pessoais do tipo “paraler” ou “gostei” são frequentes nesses sistemas, sendo que essas *tags* são feitas para uso exclusivo e, assim, podem prejudicar a recuperação de conteúdos.
- **Marcações ambíguas:** O usuário ao fazer diversas marcações em um sistema, dependendo da quantidade de *tags* marcadas em um objeto, pode gerar muita ocorrência de palavras sinônimas, prejudicando, desta forma, a eficiência das buscas dos objetos em sua nuvem de *tags*. Tal busca se faz ineficiente devido a confusão que pode causar ao usuário. Por exemplo, um usuário pode marcar um conteúdo com a *tag* “game” e posteriormente poderá marcar outro conteúdo como “jogo”.
- **Plural, singular e palavras com acentuação:** Nesses sistemas não há uma diferenciação entre as marcações com palavras no plural ou singular, e também não se pode fazer a utilização de acentuação nessas marcações. Ou seja, o usuário pode também se confundir se caso faça *tags* no plural e singular e fica restrito a uma padronização de como usar as *tags* (sem espaços e acentuação).
- **Os sistemas não têm interligações:** Esses sistemas não têm interligações, ou seja, o usuário ao mudar de sistema ele terá que refazer sua personomia. Por exemplo, no Delicious ele tem sua personomia, sua organização etc, se esse mesmo usuário começar a utilizar o Flickr, ele não poderá fazer uso imediato das mesmas marcações do Delicious.
- **Sobrecarga de informação:** Um usuário com o tempo de uso de um ou mais sistemas baseados em folksonomia pode chegar a um ponto em que ele ficará no meio de tantas

tags que poderá confundir-se. Nesses sistemas há uma grande quantidade de informação circulando (e esse número aumenta cada vez mais), se essa folksonomia não for organizada, poderá atrapalhar ao invés de ajudar os usuários.

- **Diferentes marcações para um mesmo conteúdo:** Esse problema se reflete principalmente no momento em que se faz a busca de um objeto, pois já que as marcações são livres, os sistemas podem apresentar um mesmo conteúdo com diversas marcações diferentes para os usuários.

2.3.3 Considerações

De acordo com os problemas citados por Mathes (2004, online), percebe-se uma necessidade de melhorias nesses sistemas. Alguns autores como (WELLER, 2007), (GRUBER, 2007), (KIM et al, 2008), (KNERR, 2006), fizeram estudos sobre um ambiente no qual seria possível a junção dos conceitos de ontologia e folksonomia. Os estudos são relacionados a aplicações de ontologias como forma de solucionar alguns problemas que envolvem sistemas baseados em folksonomia. Por exemplo, o fato dos sistemas não terem interligações faz com que o usuário, ao iniciar a utilização de um outro sistema, tenha que fazer novamente suas marcações. Com uma ontologia é possível fazer uma codificação independente, o que permitirá o compartilhamento do conhecimento com outros sistemas. Alguns problemas como marcações ambíguas podem ser solucionados também através das inferências feitas pelas definições de uma ontologia, sendo a base destas ontologias, as *tags*.

Desta forma, com os sistemas baseados em folksonomia e com a participação dos usuários na formação e distribuição de conteúdos através de suas marcações, surge a ontologia de *tags* ou *Tagging Ontology*. Tal conceito pode ser definido como a utilização de uma ontologia que permite ao sistema aprender e fazer inferências dos relacionamentos através das marcações realizadas pelos usuários. Para um entendimento mais aprofundado sobre ontologia de *tags*, um estudo sobre ontologia se faz necessário. Uma síntese sobre tal conceito será apresentado a seguir.

2.4 Ontologia

A definição de Ontologia existe em dois contextos, o filosófico e o computacional. Segundo Silva (2004, p. 14), “na visão filosófica, Aristóteles define ontologia como parte da filosofia

que trata do ser enquanto ser”, já na visão computacional, Qiuou (apud VILAR, 2006, p. 29) define ontologia como,

Ontologia é a ciência que descreve entidades em um mundo e seus relacionamentos, combinando o entendimento humano de símbolos com a capacidade de processamento por máquinas.

Ou seja, é a demonstração e apresentação de um domínio e o compartilhamento das informações presentes nele feito automaticamente por computadores através de inferências. Entende-se por domínio o contexto ou sistema no qual as informações estarão sendo apresentadas, representadas e inseridas. Sendo que tais inferências devem ser coerentes e fáceis de entendimento pelos usuários. O processamento pelo computador é realizado através das inferências feitas dentro de um domínio de forma a tornar explícitas algumas relações entre as informações. Gruber (2009, online) cita também que uma ontologia pode servir para mostrar distinções e relações entre informações de um domínio em um alto nível de abstração, tornando assim os significados dos termos mais claros e objetivos.

De acordo com a citação de Gruber (2009, online), percebe-se que uma ontologia pode resolver alguns problemas de organização e de semântica dos dados. A partir disso Weller (2007, p. 109) cita que uma ontologia por si só é capaz de sanar os problemas de organização da informação e documentos, e

uma ontologia é feita por especialistas e deve ser usada para a compreensão da informação presente nos objetos e pode também retirar a ambiguidade do mesmo, sendo que não é somente para comunicação interpessoal, mas também para interações homem-computador.

Assim, com uma ontologia pode-se controlar e/ou remover alguns dos principais problemas de um sistema baseado em folksonomia, como por exemplo, fazer a organização e retirar a ambiguidade presente nas marcações dos objetos feito pelos usuários, sugerindo assim *tags* que sejam relevantes para o objeto.

Gruber (2007, online) cita que uma ontologia no contexto da web semântica é “uma tecnologia que permite a partilha e a manipulação da informação”, tal manipulação da informação se faz a partir de inferências realizadas nos relacionamentos dos conteúdos e de definições que fazem os relacionamentos e as informações dos objetos se tornarem explícitos.

Ao se fazer a junção entre o processamento dos computadores com a partilha de informação, tem-se uma das características principais de uma ontologia, que é permitir aos usuários a compreensão aprofundada do significado das informações e seus relacionamentos.

2.4.1 Características

Gruber (1993) mostra como desenvolver uma ontologia formal com o “propósito de compartilhar o conhecimento e fazer a interoperabilidade entre sistemas baseados no conceito de compartilhamento de informações”, ele cita uma série de características que uma ontologia deve ter, sendo as principais:

- **Clareza:** uma ontologia tem que ser objetiva, apresentar os relacionamentos e informações aos usuários de forma clara e lógica. E segundo Gruber (1993, p. 910) “todas as definições e resultados devem ser documentados na linguagem natural”.
- **Coerência:** a ontologia não pode apresentar as inferências fugindo do contexto no qual está inserida. Vilar (2006, p. 35) cita que “as definições que fazem parte da ontologia devem expressar parte da realidade”, apresentando assim inferências que sejam condizentes com os conceitos e com o domínio.
- **Extensibilidade:** uma ontologia deve ser modelada e bem projetada para que se possa adicionar novos termos e definições, sem que tal ontologia seja novamente refeita e/ou completamente modificada.
- **Codificação Independente:** uma ontologia deve ser especificada sem depender de símbolos e linguagens, ou seja, deve-se evitar a utilização de símbolos específicos de certos idiomas que possam atrapalhar as inferências e relacionamentos dos objetos que a compõem. A codificação independente permite que a ontologia seja usada por diferentes sistemas e diferentes plataformas.

Tais características devem ser consideradas e projetadas ao se fazer uma ontologia em que o foco seja o compartilhamento do conhecimento.

2.4.2 Objetivos

Mcguiness (2003, online), em seus estudos sobre ontologia, categorizou-as em dois tipos: simples e estruturadas. A seguir são apresentados mais detalhes sobre essas categorizações.

2.4.2.1 Ontologias Simples

No estudo de McGuinness (2003, online), foi citado que uma ontologia simples teria as seguintes características:

- **Vocabulário controlado:** este conceito faz com que todos os usuários, *databases*, autores, usem um mesmo vocabulário para os termos. Tais sistemas podem ter uma interface que estimule os usuários a usar os termos de forma controlada para facilitar as inferências.
- **Remoção de ambiguidade nos termos:** evita que um mesmo termo apareça em diferentes objetos de forma confusa.
- **Organização dos conteúdos e suporte para navegação:** a generalização das hierarquias e termos existentes ajuda os usuários na navegação, contribuindo para o usuário navegar melhor entre os conteúdos do sistema, através das subcategorias mostradas ao usuário a cada busca.
- **Suporte a buscas:** A expansão da *query* de busca através das categorias, aumentando assim alguns termos mais específicos para a categoria que deseja encontrar.

2.4.2.2 Ontologias Estruturadas

No mesmo estudo, McGuinness (2003, online), definiu também as aplicações que uma ontologia estruturada deve ter. Além de ter as aplicações citadas acima, acrescentou as aplicações a seguir:

- **Verificação de consistências:** faz uma verificação nos relacionamentos do domínio, analisando se esses relacionamentos e inferências são coerentes. Segundo Vilar (2006, p. 40), “esse tipo de verificação pode ocorrer a cada modificação que é realizada no domínio ou em determinados períodos ou marcos em um processo de evolução das informações de um universo”.
- **Suporte a interoperabilidade:** Diferentes vocabulários e sistemas podem usar as mesmas definições da ontologia criando assim um domínio único.
- **Inferências de informações:** em uma ontologia, as inferências das informações devem ser compreendidas pelas pessoas. Segundo Vilar (2006, p. 40) “as informações presentes em uma ontologia devem oferecer facilidade para serem consultadas e compreendidas”, inferindo assim situações que permita a extração da informação de forma clara.

- **Explorar generalizações e especializações das informações:** análise e abstração das informações de forma que possa variar o grau de detalhamento ou especificidade da informação, para que o usuário tenha uma visão completa sobre determinado objeto e seus relacionamentos.
- **Suporte a buscas avançadas:** se um usuário faz uma busca sobre um objeto, o sistema irá retornar aquele objeto, mas se o retorno não foi satisfatório, o usuário poderá explorar a hierarquia daquele objeto, levando a outras categorias relacionadas.

2.4.3 Componentes de uma ontologia

De acordo com Vilar (2006, p. 30) uma ontologia é composta por componentes que fazem com que o domínio e os relacionamentos sejam explícitos e formalmente definidos. Com a estruturação dos componentes principais de uma ontologia, pode-se fazer a representação de um domínio. Esses componentes são: classes, propriedades, relacionamentos, axiomas/regras e instâncias.

- **Classes:** As classes são as representações dos objetos, exatamente como em programação orientada a objetos, em que uma classe pode ser qualquer coisa que se queira representar. Exemplos de classes:
 - Turma.
 - Disciplina
 - Pessoa.
- **Propriedades:** Entende-se por propriedades as características dadas a cada conceito, Ou seja, os nomes e valores atribuídos às instâncias das classes. São exemplos de propriedades:
 - Uma pessoa tem uma propriedade “nome” em que o valor dessa propriedade seja “Joabe”.
 - Turma tem uma propriedade “número” cujo valor é “0457”.
- **Relacionamentos:** Através dos valores das propriedades alguns relacionamentos entre os valores e objetos se formam. Os relacionamentos faz com que as classes e objetos de uma ontologia se relacionem com as outras. São exemplos de relacionamentos:
 - Uma pessoa pertence a uma determinada turma.
 - Uma disciplina é ministrada por um professor.

- **Regras e Axiomas:** Regras são criadas com objetivo de impor limites que tornem os relacionamentos mais coerentes, de acordo com o domínio no qual o objeto está inserido. Vilar (2006, p.33) cita que tais regras tornam as definições de uma ontologia verdadeiras para a sua representação em um domínio. Essas regras são importantes em uma ontologia, pois são elas que estabelecem condicionais para dizer se uma ação é permitida ou não, se um objeto pode aceitar ou não determinados valores, ou se um objeto deve ter obrigatoriamente alguns valores etc. Essas inferências têm que ser coerentes com as regras pré-estabelecidas, o que permitirá a extração de conhecimentos relevantes sobre os objetos.
 - Exemplos de Regras:
 - Uma turma ativa deve conter uma disciplina.
 - Uma pessoa pertencente a uma turma deve ter um tipo entre “professor” ou “aluno”.
 - Exemplos de axiomas:
 - Um trabalho em grupo só pode ser realizado se e somente se houver mais de um aluno participando da atividade.
- **Instâncias:** É representado da mesma forma que se representa uma instância na programação orientada a objetos. Essas instâncias segundo Vilar (2006, p. 34) “constituem os casos pertencentes ao domínio modelado, os quais mantêm as características definidas pelas classes, bem como apresentam relações que se encontram no grupo de relações definidos pela ontologia”, sendo que essas instâncias devem ir de acordo com as regras pré-definidas.
 - São exemplos de instâncias:
 - Uma turma tem como instância, “Turma 0773”.
 - Um professor é uma instância de pessoa com nome “Thomas Gruber”.

2.5 Ontologia de Tags (*Tagging Ontology*)

Alguns autores como Gruber (2007), Kim et al (2008), Knerr(2007) fizeram um estudo sobre uma aplicação de uma ontologia baseada nas *tags* de uma folksonomia, esse conceito é chamado de “*Tagging Ontology*” ou “Ontologia de *Tags*”. Ou seja, uma ontologia em que as inferências e o aprendizado sejam baseados nas *tags* e seus relacionamentos entre objetos e usuários. Os principais objetivos de uma ontologia de *tags* são: tornar explícito o processo de

marcação, atribuir valores a essas marcações, seus relacionamentos entre usuário, objeto e *tags*, e tornar possível à interoperabilidade entre os sistemas.

A necessidade de aplicar uma ontologia baseada nas *tags* surgiu a partir da falta de formalidade que esses sistemas têm (KNERR, 2006, p. 2). Essa falta de formalidade faz com que cada sistema tenha sua forma de representar as *tags*, o que pode atrapalhar a interoperabilidade e prejudicar as inferências realizadas no domínio, pois não há um padrão a ser seguido. Segue na Figura 10 uma representação dos sistemas baseados em folksonomia sem uma ontologia de *tags* aplicada.

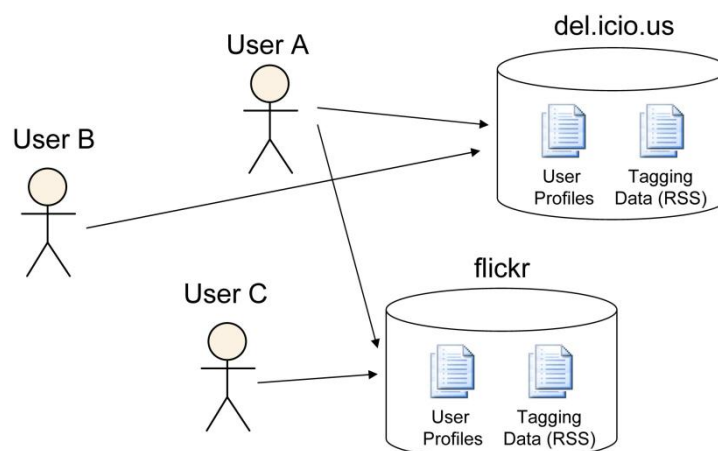


Figura 10. Falta de interoperabilidade das folksonomia atualmente. (Fonte: Knerr, 2006)

Tem-se na Figura 10 dois sistemas baseados em folksonomia (Delicious e Flickr), cada um deles guarda as *tags* dos usuários individualmente. Isso faz com que o usuário, ao mudar de sistema, defina uma nova personomia, ou seja, ele não poderá usar as marcações do outro sistema e vice-versa. Outro problema decorrente do uso de folksonomia são as ambiguidades e *tags* sinônimas, que são expressas de forma diferentes, mas possuem significados semânticos idênticos. Muitas vezes esses problemas dificultam o entendimento do contexto por parte de outros usuários, e também o processo.

Com uma ontologia baseada nas *tags* de um sistema baseado em folksonomia, é possível que os próprios usuários façam um controle de vocabulário desses sistemas através das marcações. Essa participação do usuário faz a manutenção dessa ontologia, pois a ontologia fará cada vez mais inferências sobre as *tags* e os objetos. Uma ontologia irá explicitar e tornar as marcações feitas mais coerentes de acordo com a linguagem utilizada pelos usuários, o que pode ser um fator de incentivo aos usuários no sentido de fazer marcações mais de acordo com o contexto. Essa participação ativa dos usuários, segundo

Weller (2007, p. 109) e Sharif (2007, p. 8), faz com que uma ontologia de *tags* receba um *feedback* contínuo por parte dos usuários e tenha uma evolução constante. Essa evolução se dá a partir do momento em que os usuários fazem suas marcações com novas *tags* nos diversos objetos, pois, a cada marcação, mais informações a ontologia terá e, desta forma, poderá fazer inferências mais coerentes com o objeto, o contexto e o perfil do usuário.

Com uma ontologia de *tags* também há um incentivo à interoperabilidade, como se observa na Figura 11.

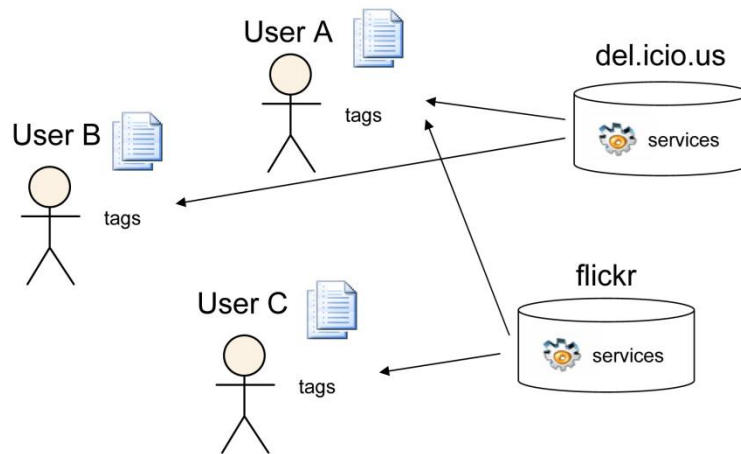


Figura 11. Interoperabilidade dos sistemas. (Fonte: Knerr, 2006)

De acordo com a Figura 11, cada usuário tem sua personomia em sistemas distintos. As personomias dos usuários podem ser compartilhadas e utilizadas em diferentes sistemas por ter sua codificação independente e, assim, proporcionar benefícios como o reaproveitamento do conhecimento produzido pelo usuário em diferentes ambientes e conseqüentemente a utilização de sua nuvem de *tags* em outros sistemas.

Gruber (2007, online) cita que,

Uma ontologia que tenha como elemento principal as marcações dos conteúdos poderá ser útil para identificar, manipular e formalizar a conceitualização da atividade de *tagging*.

Sendo assim, o conceito de *Tagging Ontology* tem como um dos objetivos a representação das entidades e atividades relacionadas ao uso das *tags*. Kim et al (2008, p. 129-130) cita as características e contribuições de uma ontologia de *tags*, que serão apresentadas a seguir.

- **Remoção da ambiguidade das marcações:** Nos sistemas baseados em folksonomia atuais, há problemas de ambiguidade das *tags*, como segue na Figura 13.

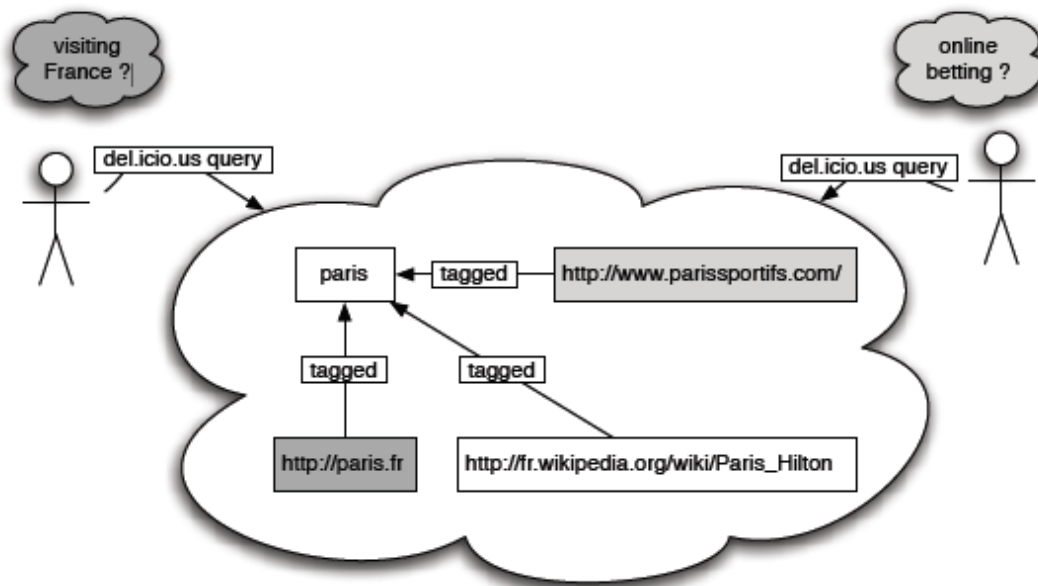


Figura 13. Ambiguidades nos sistemas atuais. (Fonte: PASSANT e LAUBLET, 2008)

De acordo com a Figura 13, um usuário faz uma marcação com o nome de “Paris” de um conteúdo que fala sobre a cidade da França, Paris. Um usuário ao buscar pela *tag* “Paris” o sistema retornaria todos os objetos marcados com tal *tag* sem analisar o contexto, e o retorno da busca faria uma exibição de objetos que referenciam “Paris Hilton”, uma loja de Paris, e a cidade. Com o processo de marcação e seus relacionamentos formados em uma ontologia de *tags*, a remoção de buscas ou marcações ambíguas podem ser diminuídas, pois a partir dos relacionamentos entre *tags*, usuário e conteúdo, a ontologia pode inferir o contexto no qual o usuário está inserido e assim retornar o que realmente ele queria buscar.

Segundo Kim et al (2008, p. 130, tradução nossa) uma ontologia baseada nas marcações dos usuários tem três características e contribuições.

- **Representação sofisticada do conhecimento:** que torna o processo de marcação e seus relacionamentos entre usuário e objeto visível e facilita a marcação dos dados.
- **Facilitação do intercâmbio de conhecimento:** uma ontologia a partir das *tags* pode ser compartilhada e usada separadamente em diferentes plataformas, ou seja, pode haver um reuso dessas *tags* por outras pessoas em outros sistemas.

- **Processamento da máquina:** ontologias no geral expõem o conhecimento humano em máquinas com intuito de fazer uma ligação automática dos dados, ou seja, fazer os relacionamentos entre os objetos de forma coerente e precisa sem ser solicitado.

Com a apresentação dos objetivos e características que compõem uma ontologia de *tags*, é necessário um estudo sobre como representar essa ontologia e quais os elementos principais. Alguns autores como Knerr (2006), Gruber (2007) e Kim et al (2008) citam os elementos principais que devem compor essa ontologia. As representações e elementos principais serão abordados na próxima seção.

2.5.2 Representação e Elementos Principais

Knerr (2006, p. 5) faz uma representação genérica dos elementos principais de uma ontologia de *tags*, sendo que tais elementos podem ser modificados dependendo da finalidade da ontologia. Na Figura 14 tem-se uma representação dos elementos padrões de uma *tagging ontology* proposta por Knerr (2006, p. 5).

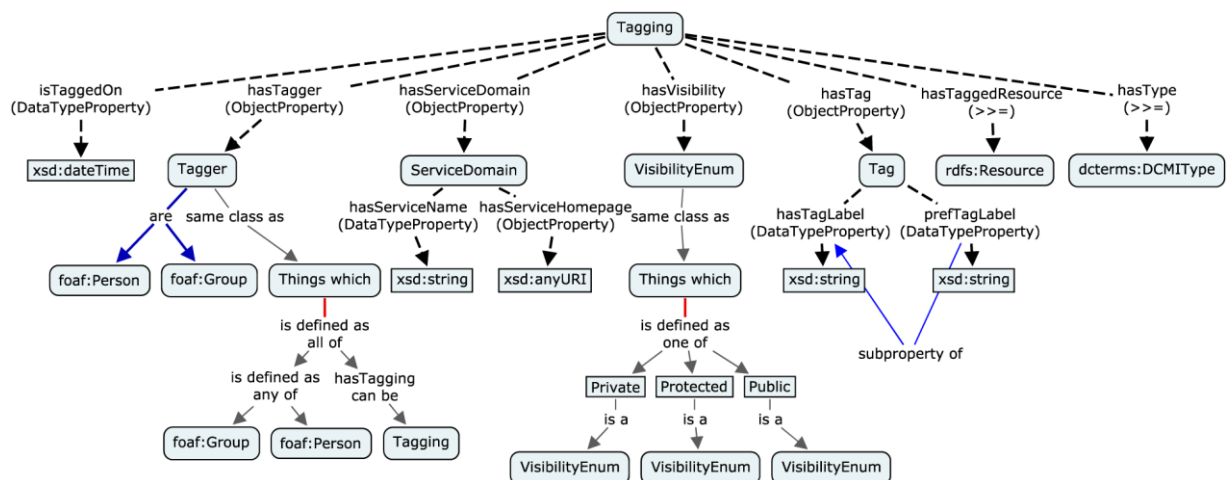


Figura 14. Ontologia de Tags de Knerr. (Fonte: KNERR, 2006)

Com o processo de *tagging* em uma folksonomia a atribuição de sentido ao objeto marcado é explicitada, pois no processo de marcação há um usuário, um conjunto de *tags* e a URL do conteúdo marcado. A Figura 14 ilustra a definição das entidades do processo de *tagging* em uma folksonomia. Assim, por exemplo, em uma rede social acadêmica um usuário faz a marcação de um objeto em uma disciplina, tal objeto passa a ser representado pela *tag* definida pelo usuário, com uma série de dados armazenados, como segue o exemplo:

- O “Usuário A” em uma rede social acadêmica marca um objeto da matéria de “Lógica de Predicados” com as *tags* “Redução ao Absurdo” e “RAA”. No processo de marcação as propriedades estudadas por Knerr são descritas a seguir:
 - Tal processo de marcação tem uma propriedade “Tempo” que representa a data em que o objeto foi marcado, para uma possível definição de relevância.
 - Há a propriedade “*Tagger*” ou “Marcador”, que pode ser um grupo, para a apresentação da folksonomia, ou um único usuário, no caso do exemplo o marcador é o “Usuário A”.
 - Na propriedade “*ServiceDomain*” ou “Domínio” é a rede social acadêmica, ou seja, o sistema no qual o usuário fez a marcação, que é representado por uma URL.
 - Há também a definição da visibilidade do objeto, se vai ser público, privado ou protegido.
 - O objeto recebe o conjunto de *tags*, que no caso são “Redução ao Absurdo” e “RAA”.
 - O objeto tem a propriedade “*Resource*” que referencia o objeto marcado. Tal referencia é feita através de uma URL.

Já na definição das entidades em uma folksonomia, Kim et al (2008, p. 130) e Knerr (2006, p. 4) fazem representações semelhantes, que são compostas por:

- **Conjunto de *Tags*:** corresponde ao grupo de todas as *tags* utilizadas.
- **Grupo do Usuário:** é a representação de onde o usuário fez a marcação, por exemplo, wikis, foruns, comunidades etc.
- **Conteúdo:** é o conteúdo marcado.
- **Popularidade da *Tag* Utilizada:** é a popularidade da *tag* representada para o conteúdo, ou seja, a frequência que uma determinada *tag* é utilizada pelos usuários para marcar objetos específicos.
- ***Tagging*:** que significa o processo de marcação dos conteúdos, representado pelo modelo de entidade de *tagging*, que é Usuário, *Tag* e Conteúdo.

Já Gruber (2007, online) faz também um representação genérica, que é representada pela Figura 15.

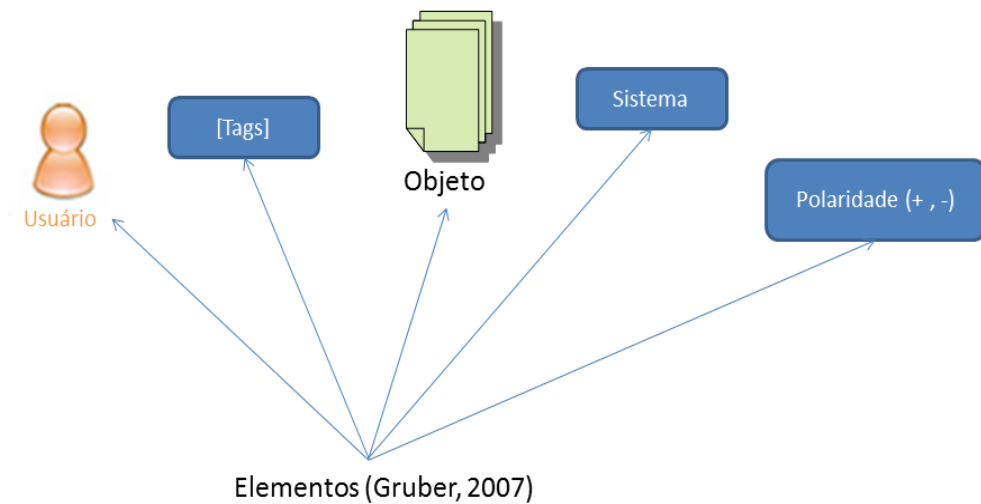


Figura 15. Representação proposta por Gruber.

Essa representação feita por Gruber (2007, online) é composta por:

- **Usuário:** o usuário que fez a marcação do objeto.
- **Conjunto de Tags:** As tags marcadas pelo usuário.
- **Objeto:** o objeto que o usuário que recebe o conjunto de *tags*, ou seja, o objeto referenciado,
- **Sistema:** o sistema que o usuário fez marcação, como por exemplo, delicious, flickr etc.
- **Polaridade:** que representa uma avaliação dos usuários para determinada *tag*. A polaridade serve para diferenciar as *tags* pessoais, como por exemplo, as *tags* “paraler”, “eu” etc, e fazer com que essas marcações não apareçam na nuvem de *tags*.

2.5.3 Aplicação em uma Rede Social

Na figura 16, tem-se uma representação da junção dos conceitos de um sistema baseado em folksonomia e com uma ontologia de *tags*.

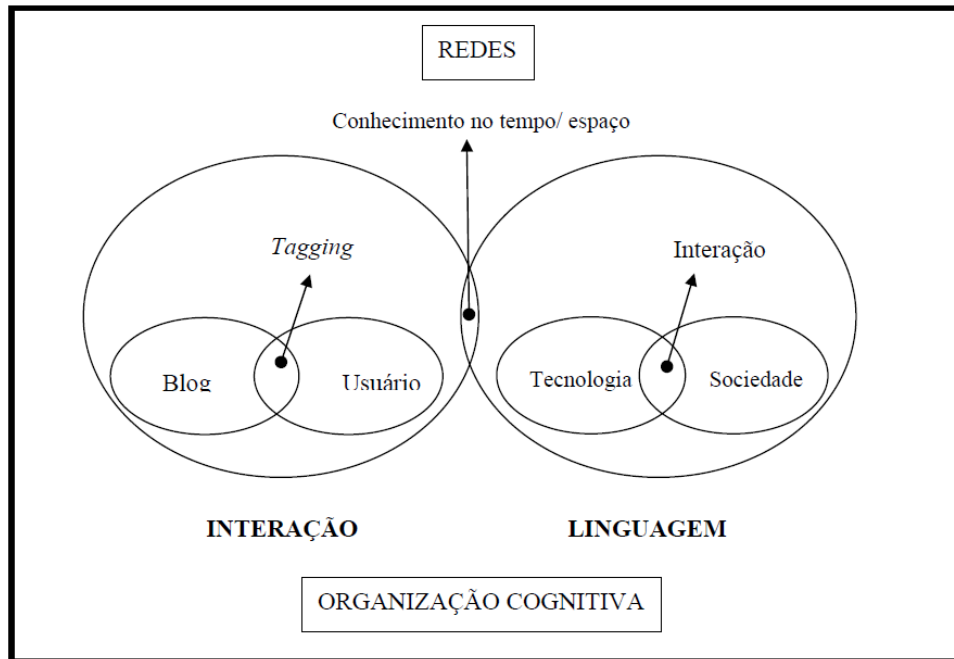


Figura 16. Junção de ontologia com Processo de Tagging. (Fonte: GONÇALVES, 2006)

De acordo com a Figura 16, percebe-se que a junção da participação dos usuários em sistemas baseados em folksonomia forma o processo de *tagging*. Essa interação e participação são formadas pela sociedade com ajuda das tecnologias disponíveis na web, a sociedade é responsável por criar documentos e *tags* de acordo com a linguagem ou cultura em que está inserido. A junção entre a interação e a linguagem utilizada forma uma organização cognitiva e um conhecimento de acordo com o tempo.

Define-se cognição como sendo a capacidade de aprender, compreender, obter conhecimento durante o tempo. Já em uma ontologia, as informações obtidas de acordo com o tempo possibilitam que sejam realizadas inferências mais coerentes sobre os conteúdos e os relacionamentos e, assim, essas inferências geram novos conhecimentos no tempo/espaço.

A geração de uma organização cognitiva nos sistemas baseados em folksonomia, tanto por parte dos usuários, quanto por parte da ontologia aplicada, possibilita uma melhoria no processo de aprendizagem de tal sistema. Dependendo de seus axiomas e regras, tal ontologia pode apresentar o conhecimento obtido para os usuários baseados no tempo de uso do sistema e de determinadas *tags* do domínio em que ele está inserido, incentivando assim, o processo de marcação através de melhores inferências sobre os significados das *tags*.

A utilização de uma ontologia de *tags* aplicada a uma folksonomia de uma rede social, dependendo dos elementos e objetivos específicos, pode diminuir ou até mesmo eliminar os problemas presentes nesses sistemas baseados em folksonomia. Com os usuários

do sistema participando das marcações, a ontologia pode fazer diversas inferências e muitas vezes incentivar os usuários a participarem cada vez mais das atividades de *tagging* do sistema. Independentemente do vocabulário utilizado, objetos marcados, o objetivo de uma ontologia de *tags* é fazer inferências coerentes através dos relacionamentos entre objetos, *tags* e usuários que existem no sistema.

3 MATERIAL E MÉTODOS

Neste capítulo será apresentada a metodologia empregada neste trabalho. O trabalho foi desenvolvido através da realização de pesquisa bibliográfica, somado à orientação prestada, essa junção permitiu o desenvolvimento e conclusão do trabalho.

3.1 Período

O período de realização compreende o 1º e 2º semestre do ano de 2011, como requisito para a disciplina de “Trabalho de Conclusão de Curso”.

3.2 Materiais

Os diferentes materiais necessários ao desenvolvimento deste trabalho foram obtidos através da biblioteca do Centro Universitário Luterano de Palmas (CEULP/ULBRA) e também através de consultas à internet. Tais materiais estão inclusos:

- Dissertações de Mestrado;
- Livros;
- Publicações Científicas;
- Artigos;
- Trabalhos de Conclusão de Curso;
- *Sites* da internet;
- Biblioteca C# *CloudLib*;

3.3 Softwares

Os softwares utilizados para a execução trabalho foram:

- Visual Studio 2010
- SQL Server 2008 R2
- Windows 7
- Notepad++

3.4 Metodologia

Neste trabalho, inicialmente, foram analisados e estudados os conceitos de folksonomia e ontologia, para utiliza-los em uma aplicação em uma rede social acadêmica. Foram estudados os conceitos de diversos tipos de redes sociais, analisando suas aplicações e objetivos.

Foi necessário também um estudo sobre os sistemas baseados em taxonomia, analisando a sua interação com os usuários e a organização desses sistemas por especialistas.

O estudo sobre folksonomia foi necessário para compreender o processo de marcação feito pelos usuários na web, entender os problemas frequentes nesses sistemas baseados em folksonomia e, por fim, aprender os conceitos necessários para a aplicação de uma ontologia de *tags* nesses sistemas de forma a sanar alguns desses problemas.

Concluído o estudo sobre folksonomia, foi desenvolvido um estudo sobre ontologia, para compreender as características, objetivos e seus componentes principais, com o objetivo de fazer estudos iniciais de uma ontologia baseada nas marcações dos usuários. Dessa forma, o estudo sobre ontologia de *tags* foi realizado com foco para aplicação em uma rede social acadêmica, para adequar a proposta do trabalho. O entendimento dos conceitos estudados foram importantes para o desenvolvimento de uma ontologia de *tags* aplicada em uma folksonomia de uma rede social acadêmica.

Para a contextualização do domínio e montagem da estrutura tanto da primeira versão do banco de dados quanto da própria rede social foram feitos alguns estudos da aplicação e algumas discussões com os professores. Nesta etapa foram definidas as tecnologias empregadas no trabalho. Após compor a estrutura do banco de dados, foram feitas diversas *views* de teste no SQL Server 2008 para que fosse possível retornar dados das *tags*.

Após a montagem e os testes, foram definidos por meio de reuniões os requisitos do sistema. Para a adaptação e correto funcionamento dos requisitos, a base de dados foi modificada, adicionando tabelas que pudessem representar o processo de marcação feito pelos usuários, e foram modificados os relacionamentos entre *tag*, conteúdo e comunidades. Com as modificações do banco de dados, foi iniciada outra bateria de testes através de novas consultas e *views*. Com as *views* testadas e com o retorno sendo satisfatório foi iniciada a implementação da biblioteca no Visual Studio 2010. Para o início da implementação, realizou-se uma nova discussão sobre a estrutura das classes e objetos da biblioteca. A partir disso, definiu-se que deveria ser criada uma única classe com diversos construtores para que fosse possível buscar os dados das consultas. O nome definido para esta classe foi “*TagResultsItem*”.

Após a implementação da classe *TagResultsItem* foram realizados testes para verificar os retornos dos dados. Com essa verificação concluída, implementou-se a classe para manipular as diversas nuvens de *tags* (por comunidade, conteúdo, usuário etc).

Para o requisito de verificação de similaridade entre comunidades foi necessário um estudo sobre o cálculo de similaridade do cosseno. Assim, foi aplicado o cálculo para que fosse possível fazer as inferências de similaridade baseado no valor retornado.

4 RESULTADOS E DISCUSSÃO

A utilização de sistemas desenvolvidos a partir do conceito da folksonomia possibilita ao usuário ter uma maior liberdade na organização de seus conteúdos, porém, tal liberdade pode acarretar em uma série de problemas, como a falta de interoperabilidade entre os sistemas, ambiguidade nas marcações etc. Com a aplicação de uma ontologia baseada nas marcações (ou ontologia de *tags*) nesses sistemas, é possível sanar alguns desses problemas. Ou seja, uma ontologia de *tags* além de poder auxiliar os usuários com as marcações, contextualizando o processo de marcação e atribuindo valor semântico a ele, pode promover também a interoperabilidade dos sistemas. Neste capítulo serão abordados os resultados desse trabalho, ou seja, uma série de etapas relacionadas à aplicação de uma ontologia de *tags* em uma rede social acadêmica, baseada nos conceitos estudados na “Revisão de Literatura”.

4.1 Contextualização do domínio

O Konnen é uma rede social desenvolvida no CEULP/ULBRA. Essa rede será utilizada em uma Instituição de Ensino Superior e terá usuários que serão representados por diversos perfis, como aluno, professor, coordenador etc., todos os perfis utilizarão *tags* para a organização dos conteúdos. Nessa rede, os usuários além interagirem normalmente como em uma rede social do estilo do *Facebook*, também podem organizar, distribuir e compartilhar os conteúdos por meio dessas marcações.

Os usuários poderão fazer marcações dos conteúdos postados pelos professores, ou qualquer outro conteúdo da Web. Segue na Figura 17 a representação da junção da folksonomia, o processo de *tagging* e ontologia de *tags* fazendo inferências para o *Konnen*.

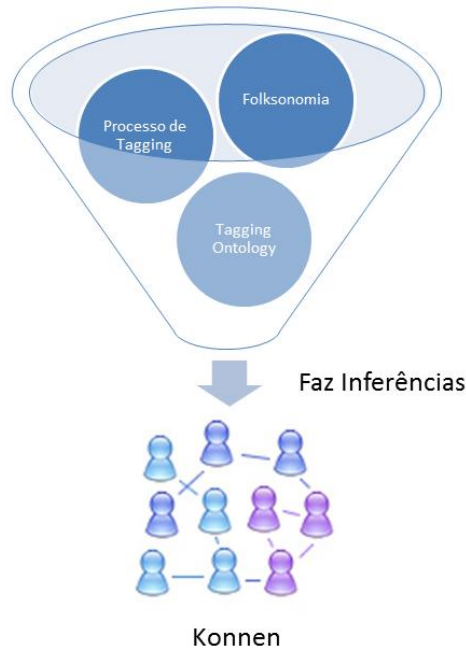


Figura 17. Folknomia + Ontologia de *Tags*

De acordo com a Figura 17, a junção dos três módulos descritos, torna possível que a ontologia de *tags* faça inferências para a rede social. A partir da junção da folksonomia e o processo de *tagging* com uma ontologia, será possível que o sistema faça inferências relevantes para os usuários dessa rede social, como, por exemplo, auxiliar no significado das marcações feitas e no compartilhamento de conteúdos através de suas nuvens de *tags*.

Assim, o módulo de folksonomia do Konnen depende da ontologia de *tags* no que tange a organização e as inferências, como sugestões de conteúdos baseados na similaridade, a contextualização das diversas nuvens de *tags* do sistema, remoção de marcações irrelevantes nas nuvens de *tags* para determinados conteúdos. Todo o processo também depende da participação ativa dos usuários, para que o sistema cada vez mais faça inferências relevantes.

4.2 Estrutura do Konnen e Visualização das Nuvens

Na rede social Konnen, as turmas são representadas pelas comunidades e cada aluno dessas turmas é alocado no início de cada semestre letivo nas comunidades que representam suas turmas. Os usuários e comunidades são os mesmos de uma estrutura chamada “Portal Ensino”, ou seja, um usuário no Portal Ensino é o mesmo usuário no *Konnen*, como também uma turma passará a ser representada por uma comunidade no *Konnen*, como segue na Figura 18.

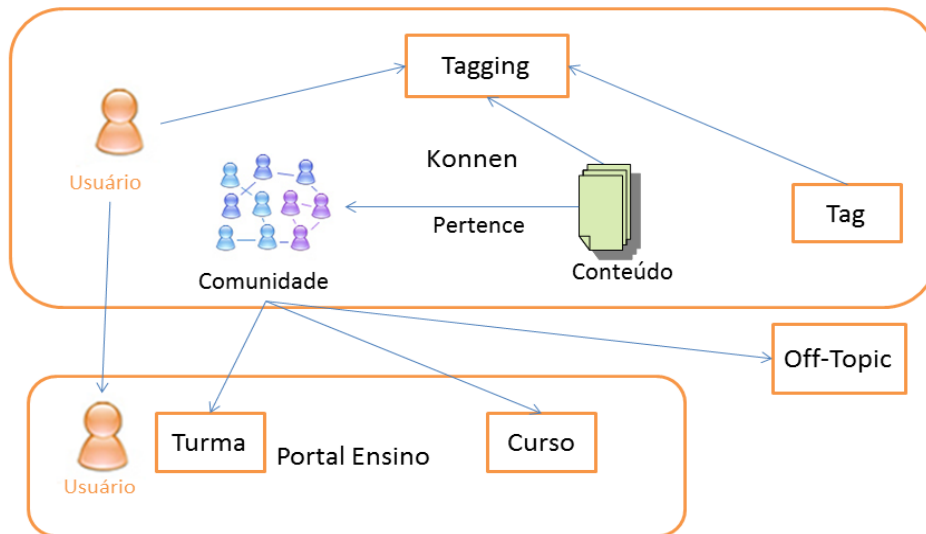


Figura 18. Estrutura do Konnen

A estrutura da rede social é formada por uma base de dados chamada “*Konnen*”. Porém, todos os dados de Turma, Curso e Usuário são buscados em uma outra base de dados chamada “Portal Ensino”. Uma comunidade pode ter dados de uma Turma ou um Curso, e os dados do Usuário são replicados para a base *Konnen*.

No sistema, o criador do objeto de aprendizagem realiza as marcações iniciais que permanecerão visíveis para os usuários, essas marcações iniciais são importantes para a divulgação e busca de conteúdo e para sugestão de *tags* para o usuário. Já as marcações feitas pelos usuários de um objeto já existente na rede serão somente visíveis para o próprio usuário e/ou para o criador do objeto, como ilustra a Figura 19.

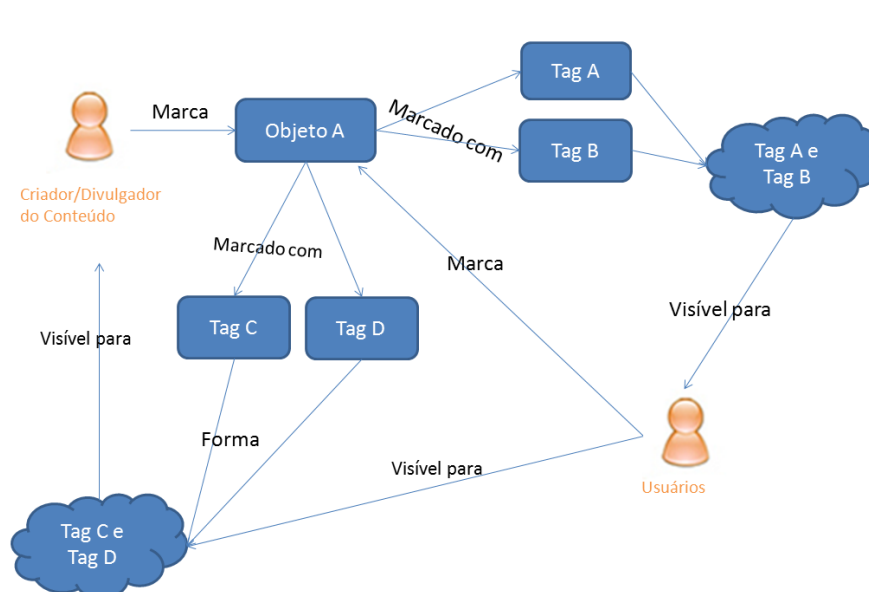


Figura 19. Representação do esquema de marcação na visão do usuário e criador do conteúdo.

Conforme apresentado na Figura 19, cada perfil (aluno, professor) que utilizar um objeto de aprendizagem (de outrem) poderá fazer suas próprias marcações e poderá também fazer a visualização das *tags* do usuário que criou o objeto inicialmente.

Os objetos postados pelos usuários são vinculados a uma disciplina e serão preponderantes para formar a nuvem de tags destas. A nuvem de *tags* que será visualizada pelos usuários do sistema é a nuvem de *tags* de cada disciplina, como ilustra a Figura 20.

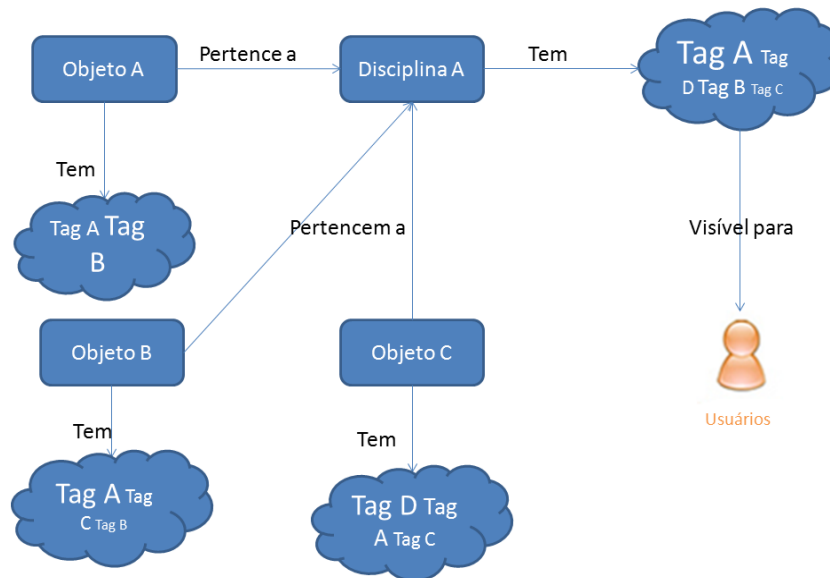


Figura 20. Visualização da Nuvem de *tags*.

Tem-se na imagem (Figura 20) a nuvem de *tags* dos objetos formando uma nuvem de uma disciplina. As nuvens de *tags* da disciplina poderão ser visualizadas pelos usuários com o perfil de aluno. Inicialmente, decidiu-se que a nuvem do objeto será usada como fonte de construção de relatórios para alguns perfis da rede (coordenador, professor etc), mas não será apresentado de uma forma geral, pois poderá causar uma sobrecarga de informação, chegando a um ponto que o usuário pode se perder em meio a um número excessivo de informações.

Uma disciplina pode ter um conjunto de nuvens de *tags*, isso porque as turmas em uma instituição de ensino são formadas por um dado período de tempo (semestre, ano) e a mudança desse período provoca uma configuração diferente no que concerne, especialmente, aos alunos. A intenção de manter esse arquivo de nuvens a cada período é preponderante para os relatórios que o sistema gerará para perfis específicos (coordenador, professor), mas todo esse aglomerado de informação não estará disponível para todos os usuários, visto que a premissa básica da marcação nos objetos de aprendizagem de cada disciplina é permitir ao

usuário (especialmente, o usuário-aluno) criar sua própria organização dos objetos das disciplinas, pois isso contribuirá com o seu processo de ensino-aprendizagem. A sobrecarga de *tags* poderá, ao invés de contribuir, prejudicar tal processo. Daí a decisão tomada de apresentação da nuvem pertinente ao período vigente.

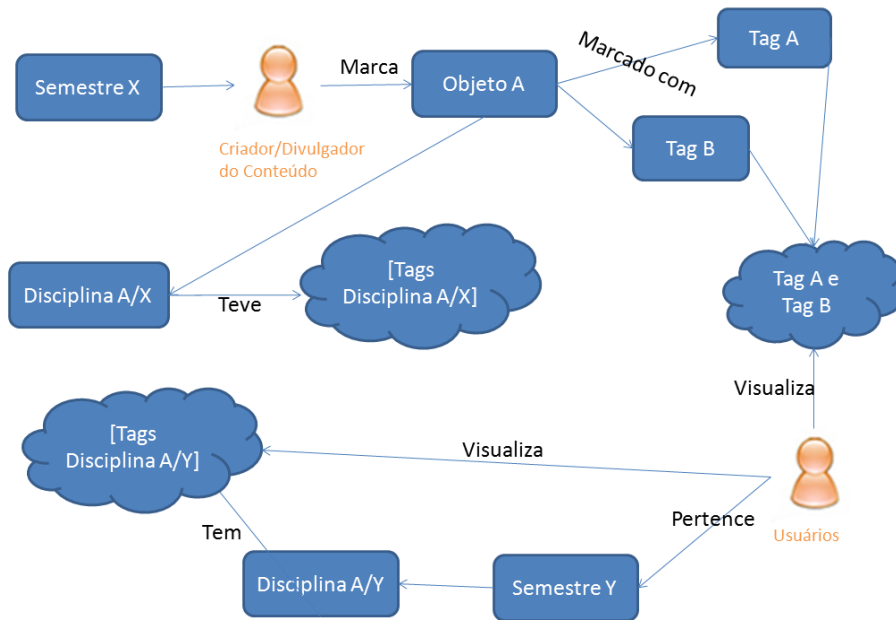


Figura 21. Tempo X Marcações de Objetos.

Na Figura 21 é apresentado o esquema explicado anteriormente, ou seja, o usuário só poderá visualizar a nuvem de *tags* de uma disciplina do semestre atual, pois, além dos fatos já elencados, a visualização de todas as nuvens da disciplina dificultaria a inferência de novas *tags* para os objetos e o objetivo inicial da marcação perderia o sentido, já que se tornaria uma mera repetição. Assim, as diferenças de marcações de um semestre para outro podem refletir desde mudança no conteúdo da disciplina, até em um perfil diferenciado da turma. Essas diferenças agregadas a outras informações, como média da turma, média de idade dos alunos, mudança de professor, entre outras, poderão se tornar indicadores do processo ensino-aprendizagem.

No que tange ao compartilhamento de conteúdos e a geração de relatórios com eficientes inferências sobre os processos de interação que ocorrem na rede, o trabalho realizado inicialmente permite estabelecimento de condicionais para a inferência de determinadas proposições. Desta forma, por exemplo, o nível de similaridade entre disciplinas de um mesmo curso foi concebido a partir de discussões com especialistas do domínio. Tais discussões contribuíram para estabelecer as nuvens de *tags* como um fator preponderante para verificar o nível de similaridade entre disciplinas de um mesmo curso ou de cursos diferentes.

Para um entendimento de como serão definidos tais condicionais, estipulou-se porcentagens para exemplificação, no entanto, tais porcentagens ainda serão motivo de estudo posterior.

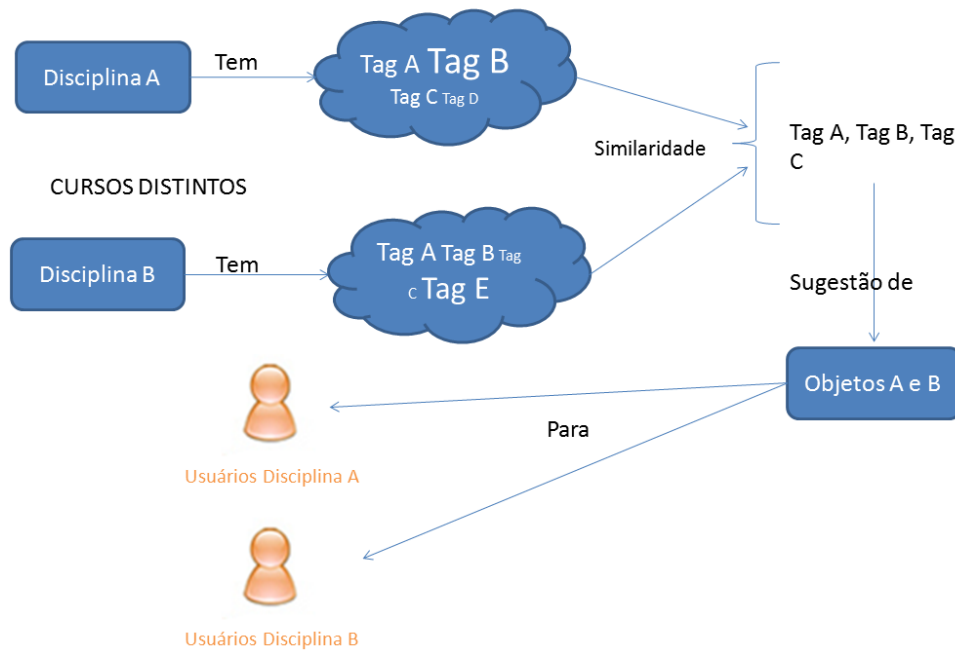


Figura 22: Sugestão por Similaridade de Tags.

Conforme apresentado na Figura 22, disciplinas de um mesmo curso com nuvens de *tags* com uma porcentagem igual ou superior a 70% de *tags* comuns são consideradas similares. Desta forma, um coordenador de curso, por exemplo, poderá verificar se a similaridade apontada pelo sistema através das marcações realizadas pelos alunos são as mesmas que existem de fato na proposta das disciplinas através de seus planos de ensino. Nas disciplinas de cursos distintos com nuvens de *tags* com tal porcentagem de tags comuns (60%), isso implica uma similaridade diferenciada, que pode acarretar em situações de interdisciplinaridade na intersecção de cursos. Para exemplificar, supõe-se que uma disciplina denominada Gestão Tecnológica I (no curso de Sistemas de Informação) que tem sua folksonomia muito similar à disciplina Economia Digital (do curso de Administração), assim alguns objetos de aprendizagem da disciplina Gestão Tecnológica I poderiam ser sugeridos aos alunos de Economia Digital e vice-versa.

4.3 Processo de Tagging no Konnen

A partir da estrutura do Konnen foi definido um conjunto de elementos que compõem uma ontologia de *tags* e o processo de *tagging* de uma rede social acadêmica. Segue na Figura 23 uma representação dos elementos dessa ontologia.

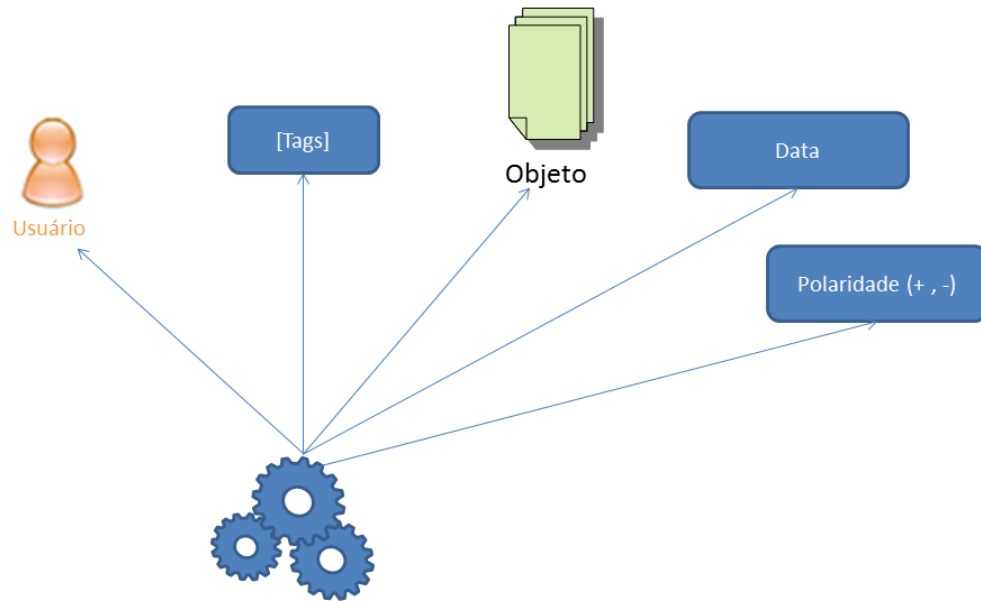


Figura 23. Elementos do processo de *tagging* para o Konnen

De acordo com a Figura 23, os elementos da ontologia são:

- usuário que faz as marcações,
- o conjunto de *tags* para um determinado conteúdo.
- o conteúdo marcado.
- data que representa o dia e hora em que foi realizada a marcação, que tem como objetivo verificar a importância das *tags* marcadas em um dado período, de forma a avaliar sua relevância. Por exemplo, no contexto do Konnen, ao avaliar as nuvens de *tags* por semestre, o professor poderá comparar o desempenho da turma com a qualidade das *tags* marcadas pelos alunos em cada período. Pode-se analisar, também, que se uma *tag* não é mais frequentemente marcada, depois de um determinado tempo ela poderá perder a relevância e ser eliminada da nuvem de *tags* através das inferências.
- Polaridade (+,-) que é útil na identificação da qualidade das *tags* para um determinado conteúdo. Ou seja, um usuário pode verificar se uma *tag* para um

conteúdo é interessante ao votar nessa *tag* como “positiva”, assim essa marcação, mesmo se tiver pouca frequência na comunidade ou em todo o sistema, poderá aparecer na nuvem de *tags* com mais destaque. Por exemplo, um usuário faz uma marcação chamada “.Net” em um conteúdo chamado “Linguagens de Programação”; o professor pode gostar dessa *tag* e marcá-la de forma positiva, aumentando sua visibilidade na nuvem.

4.4 Estrutura do Banco de Dados

A estrutura do banco de dados do *Konnen* foi adaptada para que seja possível explicitar o processo de *tagging* e fazer inferências nas nuvens de *tags* contextuais. Na Figura 24 segue a nova estrutura do banco de dados do processo de *tagging* do *Konnen*.

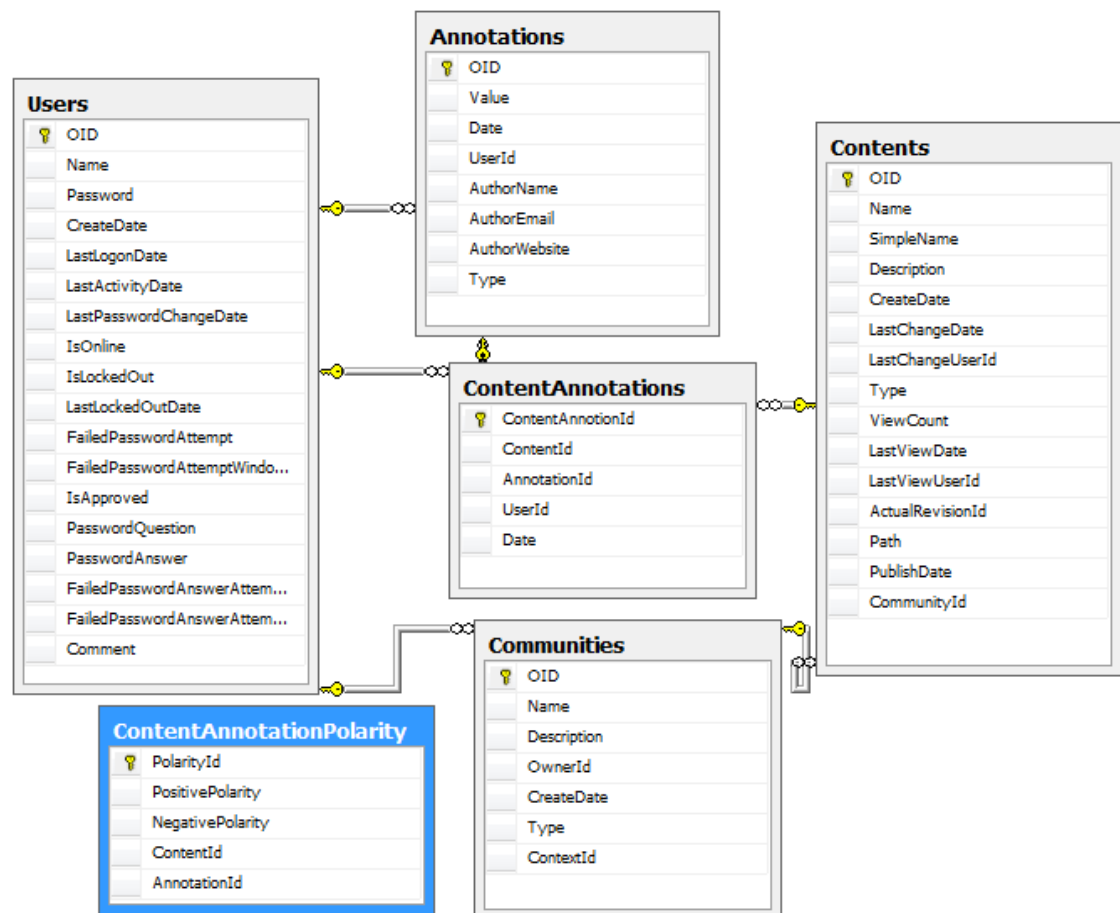


Figura 24. Estrutura do banco de dados do processo de *tagging* do *Konnen*.

De acordo com a Figura 24, tem-se a tabela *ContentAnnotations* que armazena o Id do usuário que realizou a marcação para um determinado conteúdo, o Id do conteúdo marcado, o Id da

tag utilizada e a data no qual o objeto foi marcado, representando assim o processo de *tagging*. A partir dessa tabela pode-se fazer a diferenciação de cada *tag* marcada para um conteúdo. Ou seja, o sistema pode mostrar o significado de cada *tag* para cada conteúdo. Por exemplo, o sistema pode ter várias *tags* “Paris”, mas para cada uma dessas marcações é mostrado o usuário que marcou, onde marcou e o que marcou.

Para uma estruturação de uma nuvem de *tags* com a eliminação de marcações que não são úteis para o contexto (comunidade e/ou conteúdo) foi criada a tabela *ContentAnnotationPolarity*, baseada no modelo de Gruber (2007). No modelo, cada *tag* de um dado conteúdo terá sua polaridade, ou seja, avaliações dos usuários para aquelas marcações, assim avaliações positivas tornam a marcação útil para um determinado conteúdo e resultam em uma adição no campo *PositivePolarity*. Já as marcações negativas, como *spams*, marcações pessoais etc, geram uma adição no campo *NegativePolarity*.

Essas polaridades são úteis na apresentação das nuvens de *tags* do sistema, pois se uma *tag* tem uma quantidade pequena em uma nuvem, porém, uma polaridade positiva alta, ela terá um maior destaque na nuvem. O oposto ocorre com as *tags* com polaridade negativa, ou seja, se obtiverem uma grande quantidade em uma nuvem e um valor alto na polaridade negativa, ela perderá o destaque e, em alguns casos, poderá até desaparecer da nuvem.

A tabela *ContentAnnotationPolarity* também é útil para o cálculo da similaridade entre conteúdos e comunidades. Por exemplo, uma *tag* está presente em grande quantidade em dois conteúdos, porém, em um conteúdo a polaridade está positiva, mas em outra a polaridade está negativa. Através da comparação das polaridades, apesar de ser a mesma *tag*, a relevância delas foram o oposto, a partir disso o sistema pode inferir que essas marcações, mesmo que sejam iguais, têm o valor semântico diferente para os conteúdos, logo, não são similares.

O entendimento da estrutura do banco de dados é importante para a compreensão dos requisitos do sistema, que em grande parte são baseados nos retornos das *views* do banco de dados. Os requisitos do sistema serão apresentados na seção seguinte.

4.5 Requisitos

Para início da implementação foi necessário coletar os requisitos desse sistema. Nesta seção são abordados os requisitos da rede social “*Konnen*”. Os requisitos definidos foram:

- O sistema deve retornar os dados de frequência das *tags* no formato (*tag* – objeto – quantidade):
 - *Tags* no geral;

- *Tags* conforme uma comunidade;
- *Tags* conforme um conteúdo;
- *Tags* conforme um usuário;
- *Tags* conforme um ano/semestre.
- Comparação de duas nuvens de *tags* contextuais, por exemplo, comparação entre duas nuvens de *tags* de comunidades de cursos e turmas diferentes.
- Retorno das consultas das nuvens no formato XML, para possibilitar a interoperabilidade.
- Retorno das nuvens de *tags* baseadas nas polaridades.

4.6 Aplicação

Para a aplicação dos requisitos no Konnen foi definido um conjunto de etapas. Inicialmente foram feitas as *views* no banco de dados que retornam as diversas nuvens de *tags*, seguindo assim para a implementação dos demais requisitos.

4.6.1 Views

Para os requisitos foram definidas estruturas e foram feitas consultas SQL para atender essas especificações. A Figura 25 ilustra quatro *views* que englobarão o sistema e de onde os dados das nuvens de *tags* contextuais serão retornados.

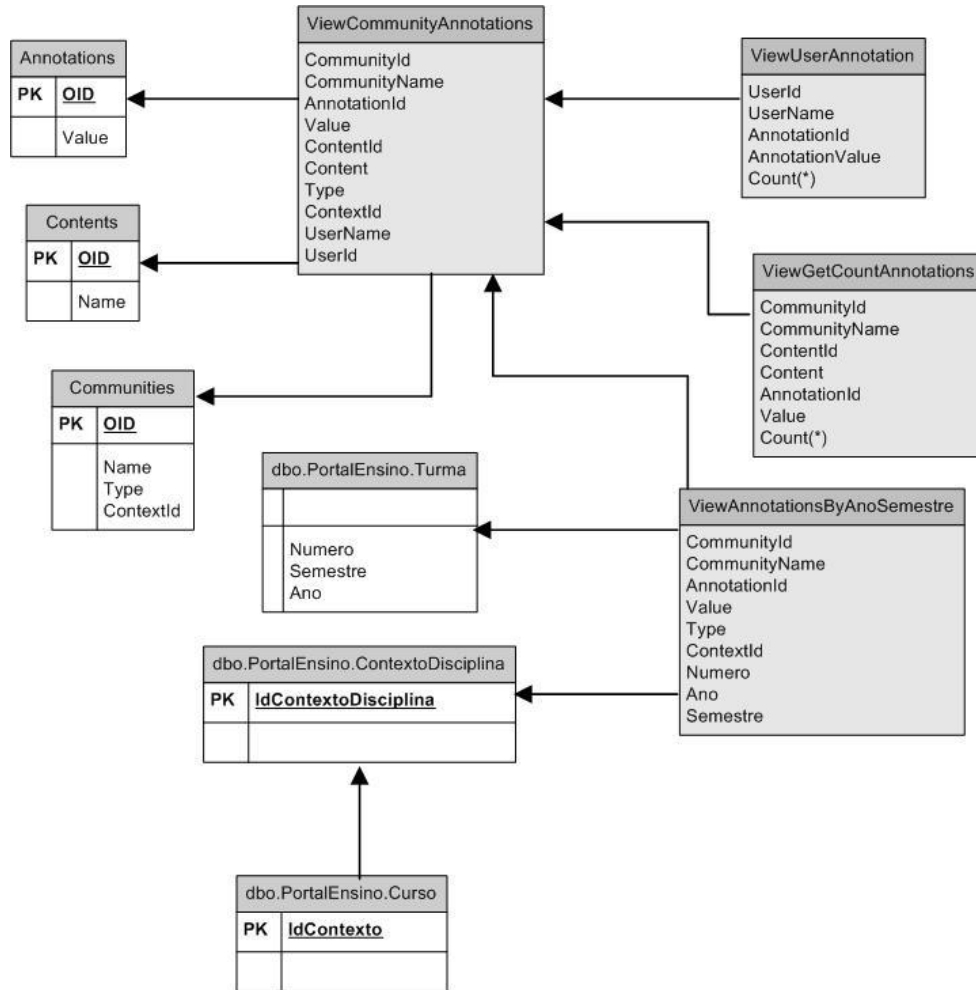


Figura 25. Estrutura das Views.

Conforme está apresentado na Figura 25, são quatro *views*:

- *ViewCommunityAnnotation* – retorna os dados do conteúdo no qual a *tag* foi utilizada, em qual comunidade foi marcada, em qual conteúdo e quem realizou a marcação. A partir dessa *view* é possível fazer o controle de visualizações dos conteúdos, pois os usuários poderão ver somente as *tags* que foram marcadas inicialmente pelo divulgador do conteúdo e as suas próprias marcações. Essa *view* também é a base para as demais consultas e *views*.
- *ViewGetCountAnnotation* – retorna os dados no formato **tag – frequência**. Essa *view* é útil para formar as nuvens de *tags* por comunidade, sem ser baseado na polaridade.
- *ViewUserAnnotation* – retorna os dados no formato **tag – frequência**. Essa *view* é útil para formar as nuvens de *tags* dos usuários. As nuvens de *tags* dos usuários não serão baseadas na polaridade, pois a nuvem de *tags* reflete a sua própria organização, não sendo necessário ser ajustada a partir da polaridade.

- *ViewAnnotationsByAnoSemestre* – uma disciplina pode ter um conjunto de nuvens de *tags*, isso porque as turmas em uma instituição de ensino são formadas em um dado período de tempo (semestre, ano) e a mudança desse período provoca uma configuração diferente no que concerne, especialmente, aos alunos. A intenção de manter esse arquivo de nuvens a cada período é preponderante para os relatórios que o sistema gerará para perfis específicos (coordenador, professor), ou seja, a visualização dessa nuvem de *tags* não poderá ser feita por alunos, pois a visualização de todas as nuvens da disciplina dificultaria a inferência de novas *tags* para os objetos e o objetivo inicial da marcação perderia o sentido, já que se tornaria uma mera repetição. Assim, as diferenças de marcações de um semestre para outro podem refletir desde mudança no conteúdo da disciplina, até em um perfil diferenciado da turma. Essa consulta retorna uma nuvem de *tags* por ano e semestre específico para atender aos requisitos de relatórios por ano e semestre. A Figura 26 ilustra a estrutura dessa *view*.

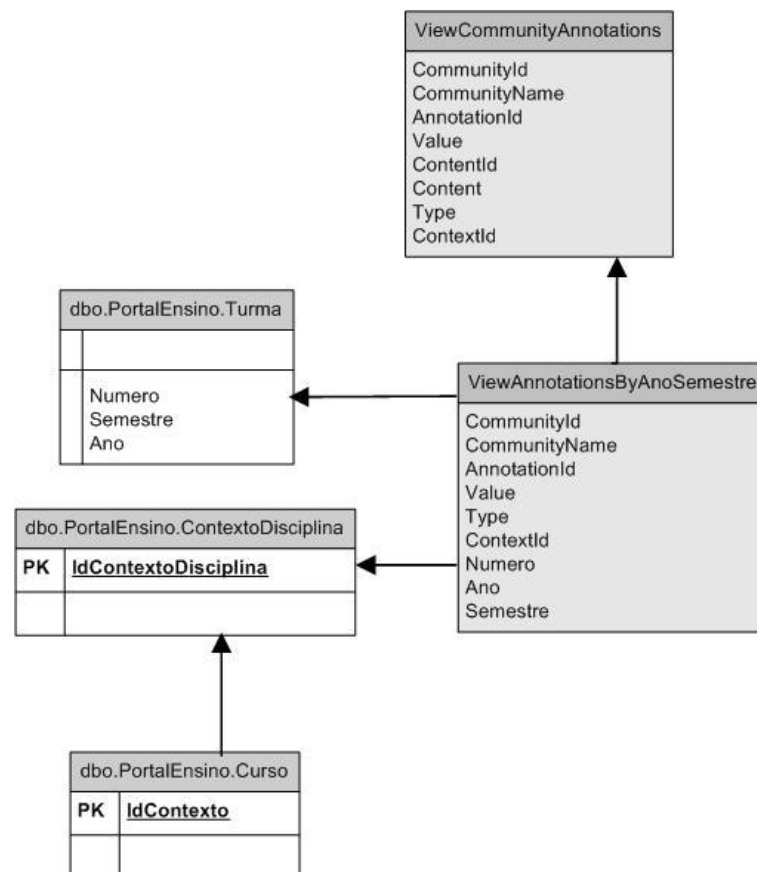


Figura 26. Estrutura *ViewAnnotationsByAnoSemestre*.

Na Figura 26 é ilustrado o processo de consulta. Essa consulta fará referência às duas bases, pois a base Konnen é necessária para obter os dados de *tags*, conteúdos e

comunidades, já a base do Portal Ensino é onde estão os dados de Turma, Ano e Semestre, que são os principais dados dessa *view*.

- *ViewGetCountAnnotationContent* – que retorna os dados no formato **tag – frequência** sendo útil para formar as nuvens de *tags* por conteúdo, sem ser baseado na polaridade.
- *ViewTagCloud* – essa *view* retorna todas as *tags* e suas quantidades, sendo possível a partir dela criar a nuvem de *tags* de todo o sistema.
- *ViewResultsPolarity* – essa *view* retorna as nuvens de *tags* de uma comunidade no formato **comunidade – tag – polaridade – frequência**, em que, o campo polaridade já é a diferença entre a polaridade positiva e polaridade negativa. Essa consulta é útil para formar a nuvem de *tags* de uma comunidade baseada nas polaridades das *tags*.
- *ViewResultsContentPolarity* – essa *view* é útil para formar a nuvem de *tags* de um conteúdo. O retorno dos dados é no formato **conteúdo – tag – polaridade – frequência**.

Todos os retornos das *views* seguem o padrão mínimo de retorno dos dados, que é *tag* e quantidade. Na Figura 27 é apresentado o retorno de dados que formam uma nuvem de *tags* simples por comunidade, ou seja, uma nuvem de *tags* que não sofre influência da polaridade.

	CommunityName	AnnotationValue	Count
1	TCC	TCC	15
2	TCC	Nuvem	14
3	TCC	Ontologia	12
4	TCC	Tags	10
5	TCC	Tagging	5
6	TCC	.net	4
7	TCC	c#	4
8	TCC	difícil	3
9	TCC	complicado	3

Figura 27. Nuvem de *Tags* simples por comunidade

De acordo com a Figura 27, os dados necessários para formar a nuvem de *tags* da comunidade “TCC” são as *tags* e suas quantidades ordenadas; o retorno da Figura 27 é o da *view* *GetCountAnnotation*. Já na Figura 28 segue um exemplo do retorno da mesma nuvem de *tags* com o cálculo da polaridade, que poderá influenciar no tamanho das *tags* na nuvem.

	CommunityName	AnnotationValue	Polarity	Count
1	TCC	TCC	-140	15
2	TCC	Nuvem	55	14
3	TCC	Ontologia	90	12
4	TCC	Tags	-53	10
5	TCC	Tagging	188	5
6	TCC	.net	0	4
7	TCC	c#	140	4
8	TCC	complicado	0	3

Figura 28. Nuvem de *Tags* por comunidade baseado na polaridade

Com a Figura 28 pode-se observar que, além de retornar a contagem das *tags* na comunidade “TCC”, é retornado também as suas polaridades gerais. A partir desses dados é possível calcular a visibilidade da *tag* na nuvem das comunidades.

Com a *view CommunityAnnotations* é retornado os dados necessários para explicitar o processo de *tagging*, ou seja, mostrar os dados detalhados do processo de marcação. Segue um exemplo dessa explicitação por usuário na Figura 29.

	UserName	CommunityName	AnnotationValue	ContentName	Date
1	konnen	TCC	complicado	Novo Conteudo	2011-11-10 00:35:28.907
2	konnen	TCC	dficil	Novo Conteudo	2011-11-10 00:35:28.903
3	konnen	TCC	Ontologia	Novo Conteudo	2011-11-10 00:35:28.887
4	konnen	TCC	Tags	Novo Conteudo	2011-11-10 00:35:28.897
5	konnen	TCC	TCC	Novo Conteudo	2011-11-10 00:35:28.877
6	konnen	TCC	.net	Novo Conteudo 2	2011-11-10 00:47:13.757
7	konnen	TCC	c#	Novo Conteudo 2	2011-11-10 00:47:13.750
8	konnen	TCC	Ontologia	Novo Conteudo 2	2011-11-10 00:47:13.733
9	konnen	TCC	Tags	Novo Conteudo 2	2011-11-10 00:47:13.740
10	konnen	TCC	TCC	Novo Conteudo 2	2011-11-10 00:47:13.720
11	konnen	TCC	Nuvem	TCC	2011-11-10 01:28:25.707
12	konnen	TCC	TCC	TCC	2011-11-10 01:28:25.700

Figura 29. Processo de *Tagging*

A partir da Figura 29 é possível explicitar o sentido da marcação, pois pode-se saber o conteúdo marcado, a *tag* utilizada e em qual comunidade o usuário marcou. Tal processo pode ser útil no momento em que houver, por exemplo, uma *tag* com nome “Manga” e um usuário faz a busca por essa *tag*, o retorno do sistema pode vir “Manga” com vários significados (manga roupa, manga desenho, manga fruta). A partir da explicitação do valor da *tag* o

usuário vai saber a comunidade que foi marcada e poder realmente saber o que estava procurando, por exemplo, comunidade com nome “Desenhos Japoneses”.

4.6.2 Implementação

A aplicação da ontologia de *tags* será feita por meio de uma biblioteca de classes na linguagem C#.net. Para a correta execução da biblioteca, é necessário que pelo menos uma tabela do banco de dados retorne a frequência das *tags*, o nome das *tags*, usuário que marcou a *tag* e a data, ou seja, uma tabela que represente o processo de marcação do sistema.

4.6.2.1 Classe *TagResultsItem*

A classe *TagResultsItem* é usada para buscar os dados para formar as diversas nuvens de *tags* e diversos retornos de dados do banco. Na Figura 30 é apresentado o diagrama da classe.

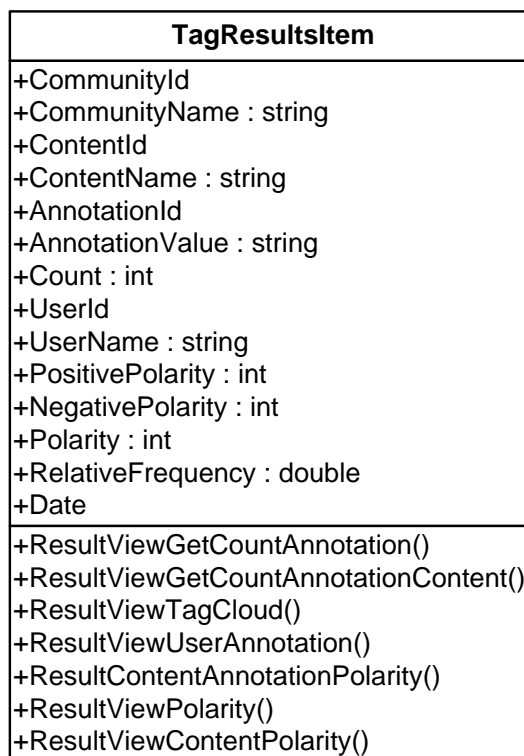


Figura 30. Classe *TagResultsItem*

De acordo com a Figura 30, a classe *TagResultsItem* tem os seguintes atributos:

- *CommunityId* – identificação da comunidade;
- *CommunityName* – Nome da comunidade;

- *ContentId* – identificação do conteúdo;
- *ContentName* – nome do conteúdo;
- *AnnotationId* – identificação da *tag*;
- *AnnotationValue* – nome da *tag*;
- *Count* – contagem das *tags* de uma determinada consulta;
- *UserId* – identificação do usuário;
- *UserName* – nome do usuário;
- *PositivePolarity* – valor das polaridades positivas de uma *tag* para um conteúdo;
- *NegativePolarity* – valor das polaridades negativas de uma *tag* para um conteúdo;
- *Polarity* – a diferença entre a polaridade positiva e negativa de uma *tag* para um conteúdo;
- *RelativeFrequency* – contagem relativa das *tags* de uma determinada consulta, ou seja, esse campo é influenciado pelo valor da polaridade
- *Date* – o tempo em que a *tag* foi marcada.

Ainda de acordo com a Figura 30 a classe tem os métodos que buscam os dados da base e armazenam na classe através dos seus diversos construtores. Esses construtores são adaptados para cada retorno de cada método listado na Figura 30. Os métodos da classe são:

- *ResultViewGetCountAnnotation* – método que busca os dados do resultado da consulta da *view ViewGetCountAnnotation*. Tal consulta é usada para formar a nuvem de *tags* por comunidade.
- *ResultViewGetCountAnnotationContent* – método que busca os dados do resultado da consulta da *view ViewGetCountAnnotationContent*. Essa consulta é usada para formar a nuvem de *tags* por conteúdo.
- *ResultViewTagCloud* – método que busca os dados do resultado da consulta da *view ViewTagCloud*. Essa consulta retorna a nuvem de *tags* geral do sistema. Os campos *AnnotationId*, *AnnotationValue*, *UserId*, *ContentId* e *Count* são utilizados no processo de *tagging* e são obrigatórios, os demais são opcionais, pois a sobrecarga de construtores resulta no preenchimento de diferentes variáveis.

Na Figura 31 segue um exemplo do construtor que busca os dados do método *ResultViewTagCloud*.

```

109 public TagResultsItem(Guid AnnotationId, string AnnotationValue, int? Count)
110 {
111     this.annotationId = AnnotationId;
112     this.annotationValue = AnnotationValue;
113     this.count = Count;
114 }

```

Figura 31. Construtor *TagResultsItem*

Na Figura 31 é ilustrado um dos construtores da classe *TagResultsItem* no qual é passado o *Id* da *tag*, o nome da *tag* e sua contagem. Esse construtor atende ao retorno da *view* que o método *ResultViewTagCloud* busca no banco. Esse método é apresentado na Figura 32.

```

202 public static List<TagResultsItem> ResultViewTagCloud()
203 {
204     KonnenEntities ke = new KonnenEntities();
205     List<TagResultsItem> listTag = new List<TagResultsItem>();
206     foreach (ViewTagCloud item in ke.ViewTagClouds.ToList())
207     {
208         listTag.Add(new TagResultsItem(item.AnnotationId, item.Value, item.Count));
209     }
210     return listTag.OrderByDescending(o => o.count).ToList();
211 }

```

Figura 32. Método *ResultViewTagCloud*

De acordo com a Figura 32, é apresentado o método que retorna uma lista de *TagResultsItem*. Da linha 206 à linha 209 são adicionados todos os itens da *ViewTagCloud*, já na linha 210 é retornado a lista ordenada pela contagem das *tags*.

Segue na Figura 33 um exemplo de outro método que retorna os dados da *ViewUserAnnotation*.

```

216 public static List<TagResultsItem> ResultViewUserAnnotation()
217 {
218     KonnenEntities ke = new KonnenEntities();
219     List<TagResultsItem> listTag = new List<TagResultsItem>();
220     foreach (ViewUserAnnotation item in ke.ViewUserAnnotations.ToList())
221     {
222         listTag.Add(new TagResultsItem(item.AnnotationId, item.Value, item.UserId, item.Name, item.Count));
223     }
224     return listTag.OrderByDescending(o => o.count).ToList();
225 }

```

Figura 33. Método *ResultViewAnnotation*

Esse método retorna os dados para a formação de uma nuvem de *tags* dos usuários, que são:

- *Id* da *tag*
- Nome da *tag*
- *Id* do usuário
- Nome do usuário

- Contagem

Para esse retorno foi criado um outro construtor, que segue na Figura 34.

```

129 public TagResultsItem(Guid annotationId, string annotationValue, Guid userId, string userName, int? count)
130 {
131     this.annotationId = annotationId;
132     this.annotationValue = annotationValue;
133     this.userId = userId;
134     this.userName = userName;
135     this.count = count;
136 }

```

Figura 34. Construtor para a *ViewUserAnnotation*.

De acordo com a Figura 34 é apresentado o construtor da classe *TagResultsItem*, que se faz útil no momento de montar o objeto com os dados retornados da *viewUserAnnotation*.

Nessa classe há outros construtores, que seguem o mesmo padrão, que servem para cada retorno de cada uma das *views* do banco de dados.

4.6.2.2 Classe TagCloud

A classe *TagCloud* manipula os dados das *views* adequando os retornos para as diversas nuvens de *tags*. Segue na Figura 35 o método que forma a nuvem de *tags* de uma comunidade.

```

30 public static List<TagResultsItem> TagCloudByCommunityId(Guid communityId)
31 {
32     List<TagResultsItem> ltc = TagResultsItem.ResultViewGetCountAnnotation();
33     List<TagResultsItem> results = ltc.Where(o => o.communityId.Equals(communityId)).ToList();
34     return results;
35 }

```

Figura 35. Método *TagCloudByCommunityId*

De acordo com a Figura 35 tem-se o método que retorna uma lista contendo os dados necessários para a formação de uma nuvem de *tags* pelo *Id* da comunidade. Na linha 32 é feita a busca dos dados da *view GetCountAnnotation* pelo método *ResultViewGetCountAnnotation* criado na classe *TagResultsItem*. A partir do retorno dos dados desse método é feita uma consulta na lista na linha 33, em que o *Id* da comunidade seja igual ao *Id* que o usuário informou. Por fim, o método retorna os dados da busca realizada na linha 33.

Para os demais métodos que formam as nuvens de *tags* também foram seguida a mesma sequência dos métodos da Figura 35, ou seja, primeiro busca no banco os resultados

das *views*, sendo retornado uma lista de *TagResultsItem*. Após o retorno é feita a consulta desejada. Na Figura 36 é apresentado um exemplo de outro método que forma a nuvem de *tags*.

```

62 public static List<TagResultsItem> TagCloudPolarityByContentId(Guid contentId)
63 {
64     List<TagResultsItem> ltc = TagResultsItem.ResultViewContentPolarity();
65     List<TagResultsItem> results = ltc.Where(o => o.contentId.Equals(contentId)).ToList();
66     return results;
67 }

```

Figura 36. Método *TagCloudPolarityByContentId*

Conforme pode ser observado na Figura 36, Na linha 64 tem-se a busca dos dados da *view ContentPolarityByContent*, na linha 65 tem-se a consulta feita na lista obtida por *Id* do conteúdo, retornando logo em seguida a lista com a consulta.

A classe *TagCloud* tem todos os métodos das possíveis nuvens de *tags* do sistema. Além dos métodos mostrados na Figura 33 e 34, os métodos restantes dessa classe são:

- *GeneralSystemTagCloud* – É o método que retorna a nuvem de *tags* geral do sistema.
- *TagCloudPolarityByCommunityId* – é o método que retorna a nuvem de *tags* de uma comunidade baseado no valor da polaridade.
- *TagCloudByContent* – é o método que retorna a nuvem de *tags* por conteúdo.
- *TagCloudByUser* – é o método que retorna a nuvem de *tags do usuário*. Essa nuvem de *tags* não sofrerá influência da polaridade, pois, como a nuvem de *tags* reflete a organização do usuário, não é necessário que o sistema faça inferências na nuvem de *tags* dele.

4.6.2.3 Classe *TagSimilarity*

A classe *TagSimilarity* tem métodos que fazem a comparação entre duas comunidades e verifica o nível de similaridade entre elas. Esses métodos são úteis para que o sistema possa fazer inferências a partir do grau de similaridade. Através dessas inferências o sistema pode sugerir aos usuários amigos, ou até mesmo os conteúdos de uma comunidade com outra, promovendo uma interdisciplinaridade.

Para o cálculo da similaridade foi utilizado a regra do cosseno. Wives (2002, p. 39) afirma que, “cada documento possui um vetor associado que é constituído por pares de elementos na forma {(palavra 1, peso 1), (palavra n, peso n)}”. Nesse caso em específico,

entende-se “documento” por “comunidade”, “palavra” por “tag” e “peso” por “frequência”. Na Figura 35 tem um exemplo de como funciona a similaridade graficamente.

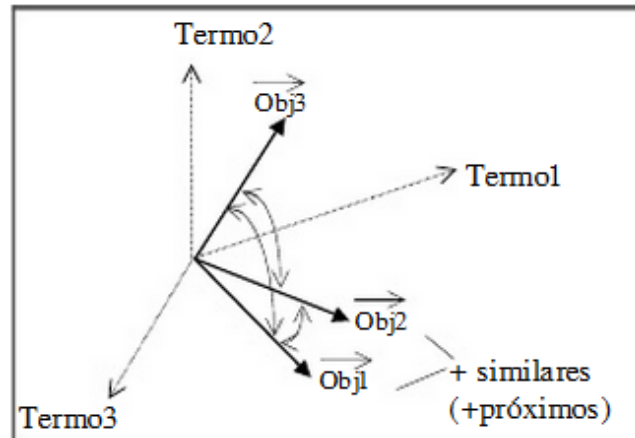


Figura 37. Gráfico da Similaridade (Fonte: Wives, 2002, p. 39)

De acordo com a Figura 37, o grau de similaridade entre os objetos é a distância entre eles, ou seja, comunidades com tags semelhantes são colocadas no gráfico, quanto mais semelhantes, mais próximos serão, e o cosseno do ângulo formado é o grau de similaridade. Segue abaixo a fórmula para calcular o grau de similaridade.

$$\text{Similaridade}(Q, D) = \frac{\sum_{k=1}^n w_{qk} \times w_{dk}}{\sqrt{\sum_{k=1}^n (w_{qk})^2 \times \sum_{k=1}^n (w_{dk})^2}}$$

Figura 38. Cálculo do ângulo de similaridade. Fonte: (Wives, 2002, p. 40)

O cálculo é feito pelo somatório do produto das frequências das tags dividido pela raiz quadrada do somatório do quadrado dos produtos das frequências das tags. Onde Q representa a comunidade do usuário e D a outra comunidade.

Inicialmente, para que seja possível a comparação entre as comunidades, é necessário que as listas de tags das comunidades tenham o mesmo tamanho e estejam ordenadas. Na Figura 39 segue o método que faz a ordenação da lista de tags das comunidades.

```

1 private static List<TagResultsItem> OrganizeList(List<TagResultsItem> userVca, List<TagResultsItem> otherVca)
2 {
3     bool add = false;
4     KonnenEntities ke = new KonnenEntities();
5     List<TagResultsItem> organizedList = new List<TagResultsItem>();
6     foreach (TagResultsItem vca in userVca)
7     {
8         foreach (TagResultsItem uvca in otherVca)
9         {
10            if (vca.annotationValue.Equals(uvca.annotationValue))
11            {
12                organizedList.Add(uvca);
13                add = true;
14            }
15        }
16        if (!add)
17        {
18            organizedList.Add(new TagResultsItem(vca.annotationId, vca.annotationValue, 0));
19        }
20        add = false;
21    }
22    return organizedList;
23 }

```

Figura 39. Método *OrganizeList*

De acordo com a Figura 39, na linha 1 o método tem como parâmetros uma lista de *TagResultsItem* da comunidade que o usuário faz parte e uma lista de *TagResultsItem* da comunidade no qual será feita a comparação. Da linha 6 a 20 é feita uma comparação de cada *tag* da comunidade do usuário com todas *tags* da outra comunidade. Se a *tag* da comunidade do usuário estiver presente na outra comunidade, então adiciona-se na lista ordenada a *tag* e a sua frequência na comunidade que está sendo comparada com a do usuários. Se após percorrer toda a lista, a *tag* não for encontrada, é adicionada na lista ordenada a *tag* que estava sendo buscada só que com a frequência zero. Após fazer a verificação de todas as *tags* é retornada uma nova lista ordenada (*organizedList*).

Na Figura 40 é ilustrado o método que calcula a similaridade entre as comunidades.


```

1 public static bool CommunitySimilarity(Guid userCommunity, Guid otherCommunity)
2 {
3     List<TagResultsItem> otherVca = TagCloud.TagCloudPolarityByCommunityIdSimilarity(otherCommunity).ToList();
4     List<TagResultsItem> userVca = RelativeFrequencyList(TagCloud.TagCloudPolarityByCommunityIdSimilarity(userCommunity).ToList());
5     List<TagResultsItem> organizedList = RelativeFrequencyList(OrganizeList(userVca, otherVca));
6
7     double numerator = 0;
8     double denominatorUser = 0;
9     double documentDenominator = 0;
10    double similarity = 0;
11
12    for (int i = 0; i < userVca.Count; i++)
13    {
14        numerator += userVca[i].relativeFrequency * organizedList[i].relativeFrequency;
15        denominatorUser += Math.Pow(userVca[i].relativeFrequency, 2);
16        documentDenominator += Math.Pow(organizedList[i].relativeFrequency, 2);
17    }
18    denominatorUser = Math.Sqrt(denominatorUser);
19    documentDenominator = Math.Sqrt(documentDenominator);
20    similarity = numerator / (denominatorUser * documentDenominator);
21    if (similarity >= 0.7)
22    {
23        return true;
24    }
25    else
26    {
27        return false;
28    }
29 }

```

Figura 40. Método CommunitySimilarity

De acordo com a Figura 40, tem-se um método que recebe como parâmetros duas listas de *TagResultsItem*, uma sendo do usuário e a outra o da comparação. Esse método retorna um booleano *true* caso as comunidades sejam similares, e *false* caso as comunidades não sejam similares. Da linha 3 à 5 são buscadas as *tags* das nuvens das comunidades baseadas nas polaridades, ou seja, a frequência é influenciada de acordo com o peso da polaridade, pois, por exemplo, em uma comunidade uma *tag* “A” tem o valor da polaridade 100, já na outra comunidade essa mesma *tag* tem polaridade -100, elas podem ser a mesma *tag*, mas tem o valor contextual completamente diferente para cada comunidade, sendo assim é calculado a frequência baseada nas polaridades, para que seja possível fazer a diferenciação dessas *tags*. Na linha 5 é feita a ordenação da lista, como analisado na Figura 41. Da linha 7 a 10 são instanciados os atributos que serão usados para o cálculo da similaridade.

Como o tamanho das listas são iguais é realizado na linha 12 um laço de repetição, sendo o limite desse laço a contagem total das *tags* da lista do usuário. Nesse laço é calculado o numerador (*numerator*). Esse numerador é acumulado em cada laço e é o resultado do produto das frequências baseadas na polaridade das *tags* analisadas. Nas linhas 18 e 19 é

realizada a raiz quadrada do somatório do quadrado das frequências das listas e, em seguida, na linha 20 é calculada a similaridade de acordo com a fórmula mostrada na Figura 40.

Foi definido inicialmente que se o resultado da similaridade for maior ou igual a 70% das *tags*, então o sistema pode inferir que essas comunidade são similares.

4.6.3 Nuvens de Tags

Com os métodos das classes da seção 4.6.2 é possível extrair os dados do banco de dados e formar uma nuvem de *tags* de acordo com o contexto, por exemplo, nuvem de *tags* geral do sistema, nuvem de *tags* por usuário, comunidade e conteúdo.

Uma mesma nuvem de *tags* pode apresentar resultados distintos, por exemplo, uma nuvem de *tags* que não tenha a contagem das marcações baseadas na polaridade pode ter um resultado diverso de quando a polaridade é considerada. Na Figura 41 é apresentada uma nuvem de *tags* simples por comunidade, tendo como base somente a contagem das *tags*.



Figura 41. Nuvem de *tags* por comunidade

De acordo com Figura 41, o entendimento da nuvem de *tags* é que as marcações com maior tamanho de fonte são aquelas *tags* que obtiveram uma contagem maior em relação às outras menores. Essa nuvem de *tags* pode ter os seus resultados modificados dependendo das polaridades das *tags*, como segue na Figura 42.

	CommunityName	AnnotationValue	Polarity	Count
1	TCC	TCC	-140	15
2	TCC	Nuvem	55	14
3	TCC	Ontologia	90	12
4	TCC	Tags	-53	10
5	TCC	Tagging	188	5
6	TCC	.net	0	4
7	TCC	c#	140	4
8	TCC	complicado	0	3



Figura 42. Nuvem de tags por comunidade baseada na polaridade

A nuvem de *tags* ilustrada na Figura 42 é formada pela classe *TagCloud*, no qual busca o resultados das diversas nuvens de *tags*. Na nuvem de *tags* da Figura 42 é formada baseada nas polaridades. Pode-se observar que na nuvem ilustrada na Figura 41 as *tags* “TCC” e “Tags” foram mais destacadas devido à maior incidência dessas marcações no contexto. Porém, na nuvem de *tags* da Figura 40, as mesmas *tags*, apesar de terem uma grande contagem, obtiveram uma polaridade negativa o suficiente para diminuir o seu destaque na nuvem. Já as *tags* “polaridade”, “Tagging” e “C#”, mesmo que em pequena quantidade na comunidade, obtiveram uma polaridade positiva o suficiente para aumentar o tamanho da sua fonte na visualização da nuvem.

4.6.4 Retorno XML

Um dos requisitos de uma ontologia de *tags* é a garantia da interoperabilidade entre os sistemas. Para que seja possível o uso desses dados das diversas nuvens de *tags* do sistema foi feito uma função SQL para cada nuvem presente no sistema. Essa função retorna um xml das nuvens. Segue na Figura 43 um exemplo das funções.

```

1 ALTER FUNCTION [dbo].[fnViewResultsPolarityCommunity] (@CommunityId uniqueIdentifier)
2 RETURNS XML
3 AS
4 BEGIN
5     RETURN (Select TOP 50 * FROM ViewResultsPolarity
6     Where OID = @CommunityId ORDER BY Frequency DESC, Polarity
7     FOR XML RAW('TagCloudItem'), Elements, ROOT('TagCloud'))
8 END

```

Figura 43. Função SQL *fnViewResultsPolarityCommunity*

A Figura 43 ilustra a função que retorna a nuvem de uma comunidade baseada na polaridade. Na linha 1 é criada a função passando como parâmetro o *Id* da comunidade, sendo esse dado do tipo *uniqueIdentifier*. Na linha 2 é especificado o tipo de retorno da função. A partir da linha 5 até a linha 7 é feita a consulta em que retorna da *ViewResultsPolarity* os 50 primeiros dados ordenados pela contagem das *tags* e suas respectivas polaridades.

Na linha 7 é estruturado o retorno XML. Nessa linha é especificado no *RAW* o atributo de cada linha de dados retornados, cada elemento da *view* retornado é transformado em atributo e, por fim, é definida a raiz do XML, que no caso é “*TagCloud*”.

O retorno da função ilustrada na Figura 43 é o XML representado na Figura 44.

```

1 <TagCloud>
2 <TagCloudItem>
3 <OID>B894A87C-D768-4871-92AD-DFB461101C0B</OID>
4 <CommunityName>TCC</CommunityName>
5 <AnnotationId>57564ED6-82ED-48CE-9190-991070F96F8D</AnnotationId>
6 <AnnotationValue>TCC</AnnotationValue>
7 <Polarity>-140</Polarity>
8 <Frequency>15</Frequency>
9 </TagCloudItem>
10 <TagCloudItem>
11 <OID>B894A87C-D768-4871-92AD-DFB461101C0B</OID>
12 <CommunityName>TCC</CommunityName>
13 <AnnotationId>62BA2859-67DB-4165-A56D-90629B758FAC</AnnotationId>
14 <AnnotationValue>Nuvem</AnnotationValue>
15 <Polarity>55</Polarity>
16 <Frequency>14</Frequency>
17 </TagCloudItem>
18 <TagCloudItem>
26 <TagCloudItem>
34 <TagCloudItem>
42 <TagCloudItem>
50 <TagCloudItem>
58 <TagCloudItem>
66 <TagCloudItem>
74 <TagCloudItem>
82 <TagCloudItem>
90 </TagCloud>

```

Figura 44. Retorno XML da função *fnViewResultsPolarityCommunity*

Considerando a Figura 44, tem-se na linha 1 o *ROOT* da função, “*TagCloudItem*”, que representa cada elemento da nuvem de *tags*, como, por exemplo, da linha 3 à linha 8 é retornado no XML dados, como: *Id* da comunidade, nome da comunidade, *Id* da *tag*, nome da *tag*, diferença da polaridade dessa *tag* e a sua frequência.

Com os dados sendo retornados no formato XML, um outro sistema poderá ler esse retorno e apresentar uma nuvem de *tags*, garantindo a interoperabilidade. É possível a formação de uma nuvem de *tags* fora do domínio em questão, que é o *Konnen*, pois, como requisito básico para a formação de uma nuvem de *tags* é necessário que tenha a *tag* e sua respectiva contagem.

4.7 Conclusão

Com a implementação de uma ontologia de *tags* aplicada na folksonomia da rede social *Konnen*, a ontologia obteve um modelo que foi condizente com os requisitos em questão. Segue na Figura 45 o modelo dessa ontologia.

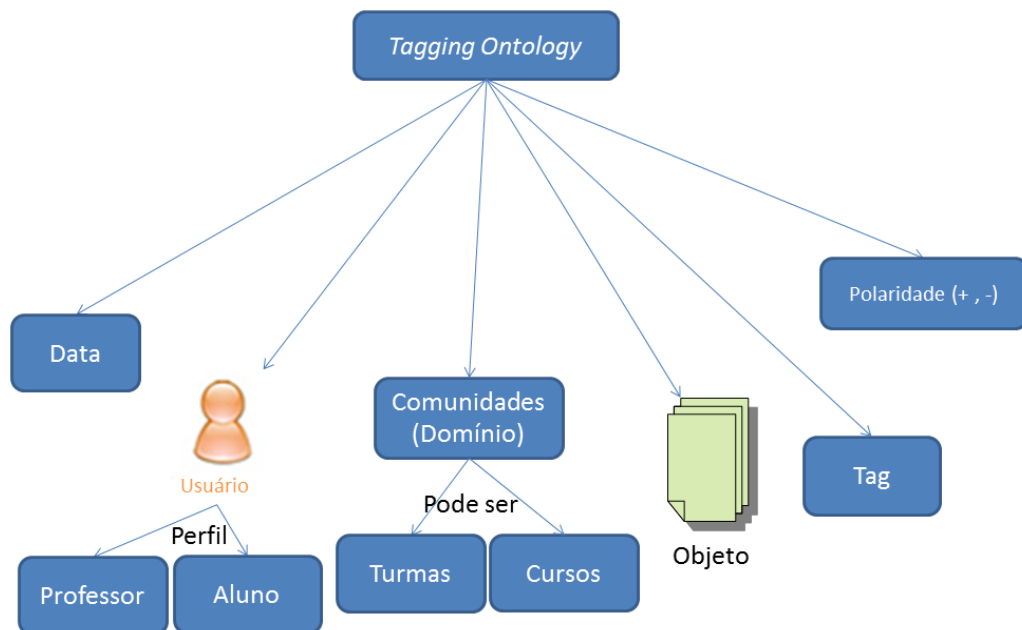


Figura 45. Ontologia de Tags para o Konnen

De acordo com a Figura 45 a ontologia de *tags* formada para o *Konnen* é composta por:

- Data – que representa o tempo que o usuário realizou sua marcação

- Usuário – que representa a pessoa que realizou a marcação e pode assumir um perfil de Professor ou de Aluno.
- Domínio/Comunidades – que representa o local onde o usuário realizou a marcação. O local da marcação pode ser em uma Turma ou em um Curso.
- Objeto – que representa o conteúdo que está sendo marcado.
- *Tag* – que representa a marcação usada no processo.
- Polaridade – que representa a avaliação dos usuários de determinada *tag* para um conteúdo.

A participação dos usuários na constituição da polaridade, através da avaliação das *tags*, possibilita um tratamento de importância de cada *tag* de forma diferenciada, pois, além de analisar a frequência que uma *tag* é marcada, também considera a importância de cada *tag* dentro do contexto. Desta forma é possível que uma *tag* com muita frequência tenha um destaque reduzido na apresentação, o que importa ao final é a relação frequência X polaridade. Esta relação permite que *tags* em comunidades diferentes possam apresentar contextos não relacionados, sendo possível gerar a similaridade entre comunidades a partir de uma análise aquém da frequência.

5 CONSIDERAÇÕES FINAIS

O trabalho teve como objetivo apresentar um estudo sobre os conceitos de redes sociais, folksonomia e ontologia, e conceitos e aplicações de *tagging ontology* ou ontologia de *tags*. Com o estudo de tais conceitos, percebe-se que, quando uma rede social tem os conceitos de folksonomia aplicados a ela, a interação e a colaboração entre os usuários tornam-se cada vez maiores, pois nessas redes sociais a participação ativa dos usuários, marcando e organizando seus conteúdos através das *tags* é um fator importante para uma melhor usabilidade.

Nos sistemas baseados em folksonomia, o usuário tem a liberdade de criar, divulgar e organizar os conteúdos. Com tal liberdade podem surgir alguns problemas, como marcações com palavra sinônimas, palavras ambíguas, questões relativas à interoperabilidade etc. Para sanar alguns dos principais problemas nesses sistemas, alguns autores realizaram estudos de como aplicar uma ontologia sem prejudicar a usabilidade e liberdade dos usuários. Surgiu assim o conceito de ontologia de *tags*, que é uma ontologia baseada nas marcações que os usuários realizam nos objetos que compõem esses sistemas.

No estudo de ontologia de *tags* foram apresentadas as principais características e os objetivos que englobam tal conceito, além de modelos genéricos dos principais elementos de uma *tagging ontology*, propostos por autores como Knerr (2006), Gruber (2007) e Kim et al (2008). Com uma ontologia aplicada a uma rede social é possível fazer inferências de novas *tags* ou apresentar uma nuvem de *tags* baseada na avaliação dos usuários da comunidade em questão.

O resultado obtido neste trabalho foi a compreensão dos conceitos de folksonomia e sua junção com uma ontologia de *tags* (*tagging ontology*) e a aplicação de modelos de interação e representação de objetos e *tags* pelos usuários. Os modelos apresentados foram preponderantes para a implementação de alguns módulos relacionados à manipulação de *tags* da rede social acadêmica Konnen, e também para a definição dos conceitos de *tagging ontology* que embasam tal sistema.

A aplicação dos conceitos de *tagging ontology* no sistema permite que os usuários visualizem diversas nuvens de *tags* contextuais, ou seja, nuvens de *tags* do domínio no qual estão inseridos. Há também a possibilidade de visualizar nuvens de *tags* baseada nas polaridades das *tags*, essas polaridades são obtidas através da participação dos usuários controlando se uma *tag* é ou não útil para determinado conteúdo.

O controle e a explicitação do processo de *tagging* permitem ao sistema mostrar o significado provável de cada *tag*, além de fazer inferências de similaridade entre as comunidades, conteúdos e usuários através desse controle no processo de *tagging*. Essas inferências na similaridade dos conteúdos do *Konnen* podem permitir aos usuários uma interdisciplinaridade, pois, haverá uma troca e compartilhamento de informações.

Na questão da possibilidade de uso das *tags* desse sistema em outro, no que concerne a interoperabilidade entre os sistemas, foi realizada com as funções SQL o retorno das nuvens de *tags* no formato XML. Esse documento XML possui dados relevantes para a formação de uma nuvem de *tags*, no formato ***tag* – frequência**.

Esses requisitos foram implementados e testados em um ambiente que não representa o ambiente do *Konnen* no servidor em produção. Portanto, ao colocá-lo no servidor em produção poderá haver modificações nos métodos, pois haverá um *feedback* dos usuários e uma maior quantidade de dados.

Em trabalhos futuros poderão ser implementados e estudados métodos para gerar relatórios baseados nas nuvens de *tags* das turmas, avaliando o seu desempenho a partir de comparações entre vários semestres. Ainda em trabalhos futuros, pode haver um estudo e implementação de um *stemming* para que seja possível a identificação de palavras sinônimas nas nuvens de *tags* de forma a agrupá-las em uma única *tag*.

6 REFERÊNCIAS

AQUINO, Maria Clara. **A Folksonomia como Hipertexto Potencializador de Memória Coletiva: Um Estudo dos Links e das Tags no delicio.us e no Flickr.** 2008. 320 p. Dissertação – Universidade Federal do Rio Grande do Sul, Porto Alegre.

COSTA, Ricardo Araújo; Oliveira, Robson Ytallo Silva; Silva, Edeilson Milhomem; Meira, Silvio Romero de Lemos. A.M.I.G.O.S: Uma plataforma para gestão de conhecimento através de redes sociais. Simpósio Brasileiro de Sistemas Colaborativos. 2008. *Anais...Washington: IEEE Computer Society*, 2008. p. 192-203. Disponível em: <[http://www.cesar.org.br/site/files/file/costa-AMIGOSPlataformaGestaoConhecimentoRedesSociais\(1\).pdf](http://www.cesar.org.br/site/files/file/costa-AMIGOSPlataformaGestaoConhecimentoRedesSociais(1).pdf)> Acesso em: 10 maio 2011.

CRUZ, Elmo Liguori. Análise de Redes Sociais. In: ____. **SNSANALYSER: Uma ferramenta para extração e análise de redes sociais a partir de comunidades existentes em sites de relacionamento.** 2008. 136p. Dissertação (Mestrado em Sistemas e Computação) - Universidade Salvador, Salvador. Disponível em: <http://tede.unifacs.br/tde_busca/arquivo.php?codArquivo=311>. Acesso em: 27 ago. 2010.

DAMME, Céline Van; HEPP, Martin; SIORPAES, Katharina. FolksOntology: An Integrated Approach for Turning Folksonomies into Ontologies. In: __ **Bridging the Gap between Semantic Web and Web 2.0.** p. 57-70, 2007. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.67.6592&rep=rep1&type=pdf>> Acesso em: 02 jun. 2011.

GOLDER, Scott A.; HUBERMAN, Bernardo A. The Structure of Collaborative Tagging System. **Journal of Information Science**, Thousand Oaks, v. 32 , n. 2, p. 198-208, 2006, Disponível em: <<http://www.hpl.hp.com/research/scl/papers/tags/tags.pdf>>. Acesso em: 09 abr. 2011.

GONÇALVES, Leandro Tavares. **Tagging: As Caravelas dos Mares Digitais.** 2006. 84 p. Monografia de Pós-Graduação em Gestão Estratégica - Universidade de São Paulo, São Paulo, São Paulo.

GRUBER, Thomas R. Towards Principles for the Design of Ontologies Used for Knowledge Sharing. **International Journal Human-Computer Studies.** 1993, v. 43, p. 907-928, Padova, Itália. Disponível em: <<http://tomgruber.org/writing/onto-design.pdf>>. Acesso em: 21 abr. 2011.

GRUBER, Thomas R. Ontology of Folksonomy: A Mash-up of Apples and Oranges. **Int'l Journal on Semantic Web & Information Systems.** v. 3, n. 1, p. 1-11. 2007. Disponível em: <<http://tomgruber.org/writing/ontology-of-folksonomy.htm>> Acesso em: 28 abr. 2011.

GRUBER, Thomas. **Ontology**. 2009, Online. Disponível em: <<http://tomgruber.org/writing/ontology-definition-2007.htm>> Acesso em: 12 jun. 2011.

GUY, Marieke; TONKIN, Emma. Folksonomie: Tidying up tags?. **D-Lib Magazine**. 2006, v. 12, n.1, online. Disponível em: <<http://www.dlib.org/dlib/january06/guy/01guy.html>>. Acesso em: 14 out. 2010.

JOHNSON, Julie E. **Managing Collective Intelligence**. França: FyP Editions, 88p. 2001. Disponível em: <http://www.axiopole.com/pdf/Managing_collective_intelligence.pdf>. Acesso em: 03 fev. 2011.

KIM, Hak Lae; SCERRI, Simon; BRESLIN, John G; DECKER, Stefan; KIM, Hong Gee. The State of the Art in Tag Ontologies: A Semantic Model for Tagging and Folksonomies. In: __ DCMI International Conference on Dublin Core and Metadata Applications. **Dublin Core Metadata Initiative** 2008. p. 128-137. 2008. Disponível em: <<http://dcpapers.dublincore.org/ojs/pubs/article/viewFile/925/921>> Acesso em: 30 maio 2010.

KNERR, Torben. Tagging Ontology - Towards a Common Ontology for Folksonomies. **Retrieve** 2006. v. 14, 2008p. Disponível em: <<http://tagont.googlecode.com/files/TagOntPaper.pdf>> Acesso em: 05 abr. 2011

MATHES, Adam. **Folksonomies – Cooperative Classification and Communication Through Shared Data**. Illinois. Online, 2004, Disponível em: <<http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>>. Acesso em: 08 jan. 2011.

MCGUINNES, Deborah L. **Ontologies Come of Age**. 2003, online. Disponível em: <[http://www.ksl.stanford.edu/people/dlm/papers/ontologies-come-of-age-mit-press-\(with-citation\).htm](http://www.ksl.stanford.edu/people/dlm/papers/ontologies-come-of-age-mit-press-(with-citation).htm)> Acesso em: 17 jun. 2011.

O'REILLY, Tim. **What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software**. Sebastopol: O'Reilly Media, 2005, Disponível em: <<http://oreilly.com/web2/archive/what-is-web-20.html>>. Acesso em: 02 jun. 2011.

PASSANT, Alexandre; LAUBLET, Philippe. Meaning Of A Tag: A Collaborative Approach to Bridge the Gap Between Tagging and Linked Data. Evolution. **Citeseer**. 2008. Disponível em: <<http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-369/paper22.pdf>>. Acesso em: 04 abr. 2011.

PEREIRA, Roberto. **Folkauthority: A Aplicação do Conceito de Autoridade Cognitiva Por Meio de Folksonomia**. 2008. 161p. Dissertação de Pós-Graduação em Ciência da Computação – Universidade Estadual de Maringá, Maringá.

RECUERO, Raquel da Cunha. **Comunidades em Redes Sociais na Internet: Proposta de Tipologia Baseada no Fotolog.com**. 2006. 334p. Tese (Doutorado em Comunicação e Informação) - Universidade Federal do Rio Grande do Sul, Porto Alegre. Disponível em: <<http://www.lume.ufrgs.br/bitstream/handle/10183/8614/000582681.pdf?sequence=1>> Acesso em: 23 ago. 2010.

ROOS, Dave. How Social Networks Work. **How Stuff Works**, Atlanta. Disponível em: <<http://communication.howstuffworks.com/how-social-networks-work.htm>> Acesso em: 11 maio 2011.

SHARIF, Atefeh. Combining ontology and folksonomy: An integrated Approach to Knowledge Representation. **IFLA**, Emerging Trends in Technology Libraries Between Web 2.0 Semantic web and research technology, p. 1-13, 2007. Disponível em: <<http://www.ifla2009satelliteflorence.it/meeting3/program/assets/AtefehSharif.pdf>> Acesso em: 27 maio 2011

SILVA, Adelina. Mundos Reais, Mundos Virtuais. Os Jovens nas Salas de Chat. **La CiberSociedad**, n. 6, 2005. Disponível em: <<http://www.cibersociedad.net/textos/articulo.php?art=37>> Acesso em: 14 set. 2010.

SILVA, José Valderlei da. **Gerenciamento do Vocabulário do Usuário em Sistemas Baseados em Tagging**. 2009. 116p. Dissertação de Pós-Graduação em Ciências da Computação – Universidade Estadual de Maringá, Maringá.

SILVA, Nalva Neila Alves da. **Definição e Integração de Ontologias Utilizando a Linguagem OWL e Conduzida por um Modelo de Desenvolvimento Proposto**. 2004. 91 p. Trabalho de Conclusão de Curso (Sistemas de Informação) – Centro Universitário Luterano de Palmas, Palmas, Tocantins.

TOMAÉL, Maria Inês; ALCARÁ, Adriana Rosecler; DI CHIARA, Ivone Guerreiro. Das Redes Sociais à Inovação. **Revista Ciência da Informação**, Brasília, v. 34, n. 2, p. 93-104, maio/ago. 2005. Disponível em: <<http://www.scielo.br/pdf/ci/v34n2/28559.pdf>> Acesso em: 28 ago. 2010.

VILAR, Bruno S. C. Medonça. **Um Framework para o Desenvolvimento de Sistemas Adaptativos a Partir de Ontologias**. 2006, 147 p. Dissertação (Trabalho de Conclusão de Curso em Sistemas de Informação) – Centro Universitário Luterano de Palmas, Palmas.

WAL, Thomas Vander; **Folksonomy Definition and Wikipedia**. 2005, online. Disponível em: <<http://www.vanderwal.net/random/entrysel.php?blog=1750>>. Acesso em: 27 abr. 2011.

WELLER, Katrin; Folksonomies and Ontologies: two new players in indexing and knowledge representation. In: **Applying Web 2.0 Innovation, Impact and Implementation: Online Information 2007 Conference Proceedings** London. p. 108-115, 2007. Disponível em: <http://www.phil-fak.uni-duesseldorf.de/infowiss/admin/public_dateien/files/35/1197280560weller009p.pdf> Acesso em: 01 jun. 2011.

WIVES, Leandro Krug. **Tecnologias de Descoberta de Conhecimento em Texto Aplicadas à Inteligência Competitiva**. 2002.116p. Tese (Doutorado em Computação) – Universidade Federal do Rio Grande do Sul, Porto Alegre. Disponível em: <<http://www.scribd.com/doc/35627173/TECNOLOGIAS-DEDESCOBERTA-DE-CONHECIMENTO-EM-TEXTOS-APLICADAS-AINTELIGENCIA-COMPETITIVA>>. Acesso em: 26 de nov. 2011.