

**FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO**



**Universidade do Porto**

---

**Faculdade de Engenharia**

**FEUP**

**ESTRATÉGIAS DE CLASSIFICAÇÃO DE INFORMAÇÃO EM  
INSTITUIÇÕES DE I&D: um estudo de caso sobre o INESC-Porto**

**Bruno Eduardo Paulino**  
*Graduado em Biblioteconomia*

Dissertação submetida para satisfação parcial dos requisitos do  
grau de mestre em Gestão da Informação

Dissertação realizada sob a orientação do  
Professor Doutor António Lucas Soares,  
do Departamento de Engenharia Informática  
da Faculdade de Engenharia da Universidade do Porto

Porto, Dezembro de 2008



## RESUMO

A organização do conhecimento é um tema que tem sido objecto de estudo há vários anos. Actualmente, com os avanços tecnológicos e com o aparecimento da Internet, inovar neste campo de actuação tem-se tornado um desafio cada vez maior. Neste contexto, como intrinsecamente sabemos, o sucesso das organizações depende cada vez mais da informação e também da sua boa gestão.

Os primeiros sistemas de organização do conhecimento surgiram no Século XIX, e desde então são utilizados nos mais diversos contextos, desde as simples bibliotecas até às grandes organizações. É no entanto, a partir desses sistemas e do resultado das necessidades entretanto emergidas com o tempo, que assistimos ao aparecimento de novos tipos de sistemas, melhor adaptáveis noutros contextos, como por exemplo a taxonomia e a folksonomia.

Partindo desta perspectiva, com o objectivo de melhorar a eficiência e eficácia na gestão de informação em instituições de Investigação e Desenvolvimento (I&D), o presente estudo incidiu, para além de uma revisão de literatura relativamente aos sistemas de organização do conhecimento e um confronto entre as suas características, sobre um estudo de caso que visou analisar o comportamento dos profissionais que actuam, directa ou indirectamente no INESC-Porto no que respeita à gestão de informação. Mais concretamente estudou-se o comportamento de selecção, organização, classificação e partilha de informação desses profissionais e, a partir desse ponto, delineou-se uma estratégia de classificação de informação segundo uma combinação de sistemas de organização do conhecimento formais e informais (taxonomia e folksonomia).

Como resultado deste estudo foi possível formular a opinião de que em ambientes de I&D é extremamente importante, para a optimização da gestão de informação, o recurso a um sistema que possa incorporar canais formais e informais. Contudo, apesar dos resultados obtidos neste trabalho de investigação, o presente tema é ainda susceptível de novas abordagens.

**Palavras-chave:** sistemas de organização do conhecimento; sistemas de classificação; folksonomia; folksonomia empresarial; gestão de informação; comportamento de pesquisa de informação.

## **ABSTRACT**

Knowledge management is a topic that has been a study subject for many years. Nowadays, with the technological advances and the appearance of the Internet, it represents a major challenge to innovate in this area. In this scope, as we can realize, the organization's success depends even more on information and its reliable management.

The first management organization systems arose in the 19th Century and since then have been used in the most various contexts, starting in the small libraries and reaching bigger organizations. It is however from these systems and from the needs emerged along the years that we are able to testify the appearance of new types of systems, capable of a better adjustment to different scenarios, as it is the particular example of the taxonomy and folksonomy.

Starting from here, and aiming to improve the efficiency and efficacy in information management in the Research and Development Institutions (R&D), this study focused on, besides a literature approach regarding knowledge management systems, a case study aiming to portray the behaviour of the professionals operating direct or indirectly at the INESC – Porto, in what the information management concerns. More specifically, a study was made in order to analyse the selection, organization, classification and information share behaviour and, from this stage, a strategy was designed on the information classification, following a combination of formal and informal knowledge management organization (taxonomy and folksonomy)

As a result of this work it was possible to deliver the opinion that in R&D environments it is of great importance in management information optimization, to rely on a system that can assemble formal and informal channels. However, besides the obtained results from this research study, the present subject is yet an issue that requires new approaches.

**Key-words:** knowledge organization systems; classification systems; folksonomy; enterprise folksonomy; information management; information seeking behavior.

## **AGRADECIMENTOS**

No final deste trabalho, que marca uma importante fase na minha vida, cumpre-me agradecer a todos que fizeram parte deste percurso e que directa ou indirectamente em muito contribuíram para que hoje pudesse realizar este projecto pessoal.

Nada disto teria sido possível sem o Programa AlBan de Bolsas de Estudo de alto nível destinado à América Latina. Foi ele o passaporte para Portugal e para a viragem académica, profissional e pessoal. Agradeço a todos os que nele trabalham, responsável pelos imensuráveis benefícios a tantos estudantes latino-americanos.

Agradecimentos devidos ao meu orientador, Professor Dr. António Lucas Soares, pois sem o seu empenho, orientação e apoio científico não teria sido possível concretizar este projecto.

A todos os que participaram na elaboração da tese, através da sua colaboração na resposta aos inquéritos um profundo agradecimento.

A toda a minha família, em especial aos meus pais, irmãos, sobrinhos e cunhada, peças basilares na minha existência, pelo exemplo e força que sempre me transmitiram. Mesmo a milhares de quilómetros de distância foi sempre possível ouvir as suas palavras de conforto e motivação. À restante e numerosa família uma palavra pelo carinho imenso que sempre me deram e pelo inegável apoio nesta jornada.

A todos os meus amigos brasileiros a viver e a estudar em Portugal, sobretudo Karla, José Luiz e Jeann, que me ajudaram a encurtar a distância do meu país e que são hoje uma verdadeira família. Aos amigos portugueses, que me permitiram conquistar um lugar em Portugal e que foram fundamentais para o meu sucesso. Para o Jorge Flávio, Joana, Vítor Santos, Susana e Maria, um obrigado sentido e o reconhecimento merecido.

A todos dedico também este projecto. Aos que sempre estiveram mais próximo, uma palavra de amizade, reconhecimento e profundo agradecimento. Fizeram e farão sempre parte da história da minha vida.



# SUMÁRIO

RESUMO .....	iii
ABSTRACT .....	iv
AGRADECIMENTOS .....	v
SUMÁRIO.....	vii
ÍNDICE DE FIGURAS .....	ix
ÍNDICE DE TABELAS.....	x
<b>1 INTRODUÇÃO.....</b>	<b>1</b>
1.1 Motivação .....	2
1.2 Objectivos.....	3
1.3 Estrutura da Dissertação.....	3
<b>2 SISTEMAS TRADICIONAIS DE ORGANIZAÇÃO DO CONHECIMENTO .....</b>	<b>5</b>
2.1 Esquemas de Classificação.....	6
2.1.1 Estruturas de Classificação.....	6
2.1.2 Tipologia dos Esquemas de Classificação .....	14
2.1.2.1 Esquema de Classificação Universal .....	15
2.1.2.2 Esquema de Classificação Nacional.....	15
2.1.2.3 Esquema de Classificação por Assunto Específico .....	16
2.1.2.4 Esquema de Classificação Personalizado.....	17
2.1.3 Estruturas de Classificação utilizadas na <i>Web</i> .....	17
2.2 Alguns Sistemas de Organização do Conhecimento .....	20
2.2.1 Vocabulário Controlado .....	20
2.2.2 Cabeçalho de Assuntos .....	20
2.2.3 Tesouros.....	21
2.3 Ontologias.....	21
2.4 Taxonomias .....	22
<b>3 FOLKSONOMIA: SISTEMA INFORMAL DE ORGANIZAÇÃO DO CONHECIMENTO.....</b>	<b>25</b>
3.1 Características das Folksonomias.....	27
3.2 Tipos de folksonomia.....	28
3.3 Aspectos Positivos e Negativos.....	30
<b>4 FOLKSONOMIA EMPRESARIAL .....</b>	<b>35</b>
4.2 Implementação de uma Folksonomia Empresarial .....	37
4.1.1 Vantagens de implementar uma folksonomia empresarial .....	40
4.1.2 Desvantagens de implementar uma folksonomia empresarial .....	42

4.2 Exemplos de sistemas de organização do conhecimento top-down e bottom-up nas empresas.....	43
4.2.1 IBM Intranet .....	43
4.2.2 IBM Dogear (Social Bookmarking) .....	44
4.2.3 BluePages+1.....	46
4.2.4 Malibu System .....	47
4.2.5 Lucent Technologies .....	48
4.2.6 AskMe Corporation .....	49
4.2.7 STARLab: The Guide .....	50
<b>5 DIFERENÇAS ENTRE OS SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO.....</b>	<b>52</b>
5.1 Alguns Estudos Anteriores.....	52
5.1.1 Tesouro vs Ontologia .....	52
5.1.2 Classificação vs Taxonomia vs Ontologia .....	53
5.1.3 Taxonomia vs Folksonomia .....	54
5.2 Folksonomia vs Sistemas Formais de Classificação .....	59
<b>6 COMPORTAMENTO DE PESQUISA DE INFORMAÇÃO DE ENGENHEIROS E CIENTISTAS.....</b>	<b>66</b>
6.1 Estudos Anteriores.....	66
6.2 O comportamento na selecção, organização, classificação e partilha de informação: um estudo sobre o INESC-Porto .....	72
6.2.1 Objectivos.....	73
6.2.2 Público-alvo .....	73
6.2.2.1 INESC-Porto .....	74
6.2.2.2 Inquiridos .....	74
6.2.3 Resultados .....	76
6.2.3.1 Organização e Partilha de Informação .....	76
6.2.3.2 Fontes de Informação.....	80
6.2.4 Conclusões .....	84
<b>7 PROPOSTA DE UMA COMBINAÇÃO DE ESTRATÉGIAS <i>TOP-DOWN</i> E <i>BOTTOM-UP</i> NA GESTÃO DE INFORMAÇÃO DO INESC-PORTO.....</b>	<b>88</b>
<b>8 CONCLUSÕES E TRABALHOS FUTUROS.....</b>	<b>95</b>
8.1 Conclusões .....	95
8.2 Trabalhos Futuros .....	97
<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>99</b>
<b>ANEXO A.....</b>	<b>105</b>



## ÍNDICE DE FIGURAS

FIGURA 1 – Modelo de estrutura hierárquica .....	7
FIGURA 2 – Modelo de estrutura em árvore .....	9
FIGURA 3 – Modelo de estrutura paradigmas .....	11
FIGURA 4 – Modelo da estrutura análise facetada .....	12
FIGURA 5 – Representação de uma tag cloud.....	28
FIGURA 6 – Representação da Broad e Narrow Folksonomy .....	29
FIGURA 7 – Representação das etapas de implementação .....	90
FIGURA 8 – Representação do esquema folksonómico .....	91
FIGURA 9 – Representação da taxonomia baseada na folksonomia .....	93

## ÍNDICE DE TABELAS

TABELA 1 – Comparativo entre serviços da IBM baseados em tags .....	44
TABELA 2 – Informações sobre os inquiridos.....	74
TABELA 3 – Organização, classificação e guarda de documentos e conteúdos .....	78
TABELA 4 – Frequência de utilização das fontes de informação .....	81
TABELA 5 – Factores que influenciam na escolha e acesso às fontes de informação electrónicas.....	82
TABELA 6 – Dificuldades encontradas na pesquisa de informação.....	83



# 1 INTRODUÇÃO

A organização do conhecimento não é uma preocupação recente e, desde há vários anos, tem sido objecto de estudo, levando especialistas a desenvolver trabalhos em sistemas e abordagens para encontrar a melhor solução para esta realidade. É no entanto com o aparecimento da *Internet*, que se assiste a um crescimento exponencial do volume de informação disponível, facto que terá levantado alguns problemas à Ciência da Informação, relacionados com a gestão desse acréscimo de conteúdos emergentes.

Em alguns contextos, como na própria *Web*, a maioria dos sistemas tradicionais para a gestão do conhecimento já não satisfazem mais as necessidades de organização e classificação de conteúdos, pois, devido à sua estrutura pouco flexível e também à sua incapacidade em incorporar rapidamente novos temas, tornaram-se sistemas pouco atractivos e aplicáveis. No entanto, e contrariamente aos anteriores, os esquemas facetados não se enquadram nestas características e, por este motivo, são sistemas bastante utilizados na *Internet*, principalmente em directórios de conteúdos.

A necessidade de crescente inovação com o objectivo de acompanhar a própria evolução das novas tecnologias de informação, nomeadamente em matéria de organização do conhecimento, originou o aparecimento de alguns sistemas e abordagens. Alguns deles estão a ser reaproveitados de outros campos de estudo, como é o caso da taxonomia, antes direccionada para a classificação de seres vivos, e actualmente utilizada também para classificação de outros conteúdos. Outros ainda direccionados para a informalidade e facilidade na utilização, para além de incentivar uma classificação colaborativa, como no caso da folksonomia. Tanto mais que a própria folksonomia é também uma forma de organizar conteúdos a baixo custo, ao contrário de uma taxonomia bem definida. Hoje pode observar-se que, no caso concreto da *Web*, a folksonomia tem sido cada vez mais adoptada e bem sucedida, pois é mantida e actualizada através da colaboração dos seus utilizadores, da mesma forma que a própria *Internet*.

Actualmente não restam dúvidas que as empresas e instituições dependem desta gestão e organização da informação para o seu sucesso. E, da mesma forma que obter e organizar informações é claramente necessário, ponderar custos, eficiência e eficácia é igualmente importante.

Segundo (Choo 1996), existem três áreas distintas em que a criação e a utilização da informação desempenham um papel estratégico na determinação da capacidade de crescimento e adaptação de uma organização: a) na tomada de decisões importantes; b) na avaliação das mudanças e desenvolvimento do seu ambiente externo; c) na criação, organização e proces-

samento de informação com o objectivo de gerar conhecimento através da aprendizagem organizacional.

No entanto, para (Davenport and Prusak 2000), a mera existência de conhecimento em algum lugar da organização não é o suficiente, pois ele torna-se uma mais valia somente se estiver acessível, e o seu valor aumenta de acordo com o seu nível de acessibilidade. Segundo os autores, é comum a duplicação de esforços na resolução dos mesmos problemas, pois, muitas vezes, as soluções desenvolvidas não são partilhadas dentro da empresa.

Como parte de uma tentativa de melhorar o acesso aos recursos informacionais via vocabulário controlado e organização do conhecimento (Tudhope and Nielsen 2006), os Sistemas de Organização do Conhecimento (SOC) são importantes ferramentas para a gestão estratégica das organizações.

Desta forma, decidir qual é ou quais são os métodos de classificação mais apropriados ao seu ambiente de trabalho é um factor importante para as organizações, pois uma ferramenta ou abordagem mal escolhida pode dificultar ou até mesmo inviabilizar o rápido acesso às informações.

## **1.1 Motivação**

Para o desenvolvimento deste estudo foram, desde logo, identificados três grandes factores que motivaram a pesquisa de informação nesta matéria: a) a formação na área da Gestão da Informação; b) o importante papel que tem a informação nos contextos empresarial e institucional; c) os actuais desafios na gestão do conhecimento.

O principal objectivo era encontrar soluções optimizadas na organização do conhecimento em empresas e instituições, através da implementação de sistemas que fossem ao mesmo tempo úteis e atractivos aos seus utilizadores. Desta forma, além dos sistemas formais de organização do conhecimento, o interesse recaiu sobre sistemas informais, tais como a folksonomia.

Para uma visão mais específica desta realidade, para além de realizar um estudo sobre os SOC, pretende-se aqui desenvolver um estudo de caso sobre um instituto de Investigação e Desenvolvimento (I&D), sendo esta uma área de actuação onde o principal motor é a própria informação.

Desta forma, a questão central do presente trabalho é saber se as folksonomias tornam mais eficazes a classificação e recuperação de informação por profissionais que actuam nesta área.

Assim, a tese aqui proposta é que uma combinação de estratégias de classificação *top-down* e *bottom-up* torna mais eficiente e eficaz a gestão de informação em instituições de I&D.

## **1.2 Objectivos**

Partindo do enquadramento anteriormente exposto, pretende-se neste estudo, melhorar a eficiência e eficácia na gestão de informação em instituições de I&D, mais especificamente no Instituto de Engenharia de Sistemas e Computadores do Porto (INESC-Porto), através da utilização de uma combinação de estratégias de classificação *top-down* (ou seja, estruturadas hierarquicamente) e *bottom-up* (sem uma estrutura hierárquica).

Sucintamente, este objectivo foi definido utilizando duas abordagens específicas:

- Estudar o comportamento de selecção, organização, classificação e partilha de informação dos profissionais que actuam, directa ou indirectamente, no INESC-Porto;
- Definir uma estratégia de classificação de informação segundo uma combinação de SOC formais e informais.

## **1.3 Estrutura da Dissertação**

Para atingir os objectivos propostos neste trabalho, o desenvolvimento do presente estudo incidiu sobre oito capítulos. No capítulo que agora se encerra, é feita uma abordagem geral do tema “gestão do conhecimento”, exposição das motivações para o desenvolvimento deste estudo, definição dos objectivos e uma esquematização da estrutura da dissertação.

No segundo capítulo, é feita uma abordagem sobre os sistemas tradicionais de organização do conhecimento. Dentro deste tema, é definido o conceito de esquemas de classificação, as suas estruturas e tipologias, como também são abordadas as estruturas de classificação utilizadas na *Web* e alguns SOC, além da ontologia e taxonomia.

No capítulo 3, é apresentada a folksonomia, um sistema informal de organização do conhecimento. É definido o seu conceito, características e tipologias existentes. Numa análise esquematizada, evidenciam-se alguns pontos positivos e negativos retirados de literatura existente.

Sequencialmente ao tema supra, no capítulo 4, é feita uma abordagem sobre a folksonomia implementada em empresas, a folksonomia empresarial. São apresentados tópicos relacionados com a implementação deste sistema e descritas as suas vantagens e desvantagens. Por último, alguns casos de empresas e instituições que utilizaram este sistema são abordados.

O capítulo 5 versa sobre a comparação entre as características dos sistemas formais e informais de organização do conhecimento. Num primeiro momento, são apresentados alguns estudos anteriores relacionados com o tema, e num segundo momento, é realizado um estudo que visa comparar a folksonomia a outros sistemas de classificação.

Com base nesta contextualização dos SOC, no capítulo 6, a ênfase é dada ao estudo do comportamento de pesquisa de informação de engenheiros e cientistas. Na sequência do estudo realizado sobre o tema, é apresentado o estudo de caso realizado no INESC-Porto, sobre o comportamento dos seus profissionais que diariamente estão envolvidos com a gestão de informação.

É pois da compilação de todas as informações retiradas até então, que, no capítulo 7, é delineada uma proposta de uma combinação de estratégias *top-down* e *bottom-up* na gestão de informação do instituto.

O trabalho é encerrado com a apresentação das conclusões do estudo e com algumas propostas para trabalhos futuros.

## 2 SISTEMAS TRADICIONAIS DE ORGANIZAÇÃO DO CONHECIMENTO

Com os avanços tecnológicos, milhares de informações tornaram-se acessíveis a partir dos mais diversos meios de comunicação. Ter controlo, fácil acesso e poder gerir de forma integrada essas informações passou a ser uma necessidade e um diferencial para as organizações.

A organização do conhecimento não é uma preocupação recente e, ao longo de centenas de anos, o esforço dos especialistas tem sido no sentido de criar abordagens apropriadas para o fazer. A partir desses esforços surgiram desde os tradicionais esquemas de classificação para as bibliotecas até aos métodos de classificação para a *Web*, como por exemplo a Folksonomia.

Relacionada com a Ciência da Informação, a organização do conhecimento é um campo de estudo que se refere a actividades como descrição, indexação, classificação de documentos e análise de assuntos. Numa perspectiva mais restrita, trata-se de SOC como registos bibliográficos, sistemas de classificação, tesouros, redes semânticas etc. (Broughton, Hansson et al. 2005). Numa visão simplista, poderá dizer-se que é uma forma de classificar o conhecimento através de técnicas e abordagens específicas.

O acto de categorizar numa classe ou classes, atribuir grupos, como por exemplo ordens e famílias, de acordo com algumas relações ou atributos comuns, pode ser definido como classificação (Wiktionary 2008). Para (Uddin, Mezbah-ul-Islam et al. 2006), a classificação é uma maneira de organizar as informações em categorias, de acordo com as suas semelhanças e diferenças. Ela pode ser usada como uma representação rica do que é conhecido e, portanto, útil na comunicação e na geração de um novo ciclo de exploração, comparação e teorização. (Kwasnik 1999), destaca que uma boa classificação reúne conceitos numa estrutura útil. Para (Tristão, Fachin et al. 2004),

"Classificação significa a acção e efeito de classificar, e classificar significa ordenar e dispor em classes. Uma classe consiste num número de elementos qualquer (objectos e ideias) que possuem alguma característica comum, pela qual devem ser diferenciados de outros elementos e, ao mesmo tempo, constitui a sua própria unidade."

O principal objectivo da classificação é ajudar indivíduos e máquinas na estruturação e localização de informações (van Rees 2003).

Uma versão simplificada dos SOC são as listas de conteúdos e índices de livros. O conhecimento está expresso em forma de texto, e o SOC é uma ferramenta suplementar que ajuda o leitor a se localizar no texto. Mas como estas ferramentas têm se tornado cada vez mais com-



plexas e com mais funções, começaram a adquirir outros nomes, como linguagens de recuperação, taxonomias, categorizações, dicionários, tesouros ou ontologias. Actualmente são vistos como sistemas que organizam, gerem e recuperam informação (Vickery 2008).

## ***2.1 Esquemas de Classificação***

Os esquemas de classificação agrupam recursos relacionados numa estrutura hierárquica ou em árvore. Cada nó dessa estrutura é designado com um código (denominado notação) que pode ser alfabético, numérico ou alfanumérico. Esses esquemas podem ser gerais (aplicam-se a todas as áreas do conhecimento) ou específicos para uma disciplina, tema particular, literatura nacional, etc. Os principais tipos de classificação são os enumerativos e os facetados. A classificação enumerativa lista todos os possíveis assuntos e as suas notações numa hierarquia, e cada trabalho que é classificado ocupa uma simples localização dentro da hierarquia. A classificação facetada define as propriedades genéricas dos assuntos, denominadas facetas, e exige que o classificador identifique, em cada classe, todas as facetas definidas e adequadas ao trabalho e que depois as combine de determinada forma (Caplan 2003).

### **2.1.1 Estruturas de Classificação**

Para (Kwasnik 1999), existem quatro estruturas principais de classificação que formam a base dos diferentes esquemas de classificação. São elas hierarquias, árvores, paradigmas e análise facetada.

#### Hierarquias

Cada objecto de classificação é um elemento de uma classe. Elas são entidades organizadas em hierarquias através de relações, onde uma super-classe pode conter um conjunto de sub-classes, definidas através de regras específicas de divisão e subdivisão. São úteis quando os limites das classes são claros e as entidades estão bem definidas (Kwasnik 1999), (Fettke and Loos 2003).

## FIGURA 1 – Modelo de estrutura hierárquica

---

Information organization

Information management

Information systems

Management information systems

---

Fonte: (Transportation 2008)

No exemplo supra, a classe do topo (“*Information organization*”) é a classe mais abrangente e descreve o domínio da classificação. Ela engloba todas as subclasses e sub-subclasses. O elemento “*management information systems*”, por exemplo, é um tipo de “*information systems*”. Este, por sua vez é um tipo de “*information management*” e assim por diante.

Genericamente, o que caracteriza uma estrutura hierárquica é o facto das subclasses serem “*um tipo de*” das super-classes.

Vantagens:

- Informações completas e abrangentes;
- Hereditariedade e economia na representação de muitos atributos complexos. Muitas informações podem ser transportadas pela estrutura hierárquica;
- Permite a inferência, ou seja, o raciocínio de evidências incompletas (por exemplo, se um paciente apresenta sintomas de conjuntivite, então é possível saber também que, como um tipo de doença dos olhos, a conjuntivite partilhará propriedades com outras doenças dos olhos);
- Permite expressar as similaridades e diferenças de uma entidade em diversos aspectos (por exemplo, o solteiro é um homem que não casou. Ele possui todas as características de um homem. O homens podem ser casados ou solteiros. No entanto, um homem solteiro pertence à categoria dos solteiros.);
- Permite um alto nível de visão e perspectiva holística, ou seja, o esquema de classificação como um todo fornece uma visualização do fenómeno que ele está a representar.

Desvantagens:

- Várias hierarquias (ex.: um cão pode ser “cão como animal”, “cão como mamífero”, “cão como animal de estimação”, “cão como companhia humana” etc.);
- Diversos critérios. Existem alguns limites práticos para que a quantidade de informações não a torne demasiada complexa, pois esta estrutura não é bem concebida para acomodar distinções feitas em dois diferentes conjuntos de critérios (por exemplo, “leões selvagens versus leões no zoológico. Eles devem ser classificados como uma mesma entidade?”);
- Falta de abrangência em determinadas áreas. Em campos emergentes, onde ainda não há claro entendimento e relacionamento sobre o assunto, esta classificação é considerada difícil e inapropriada;
- Diferença de escalas. Uma entidade pode ser classificada a partir de várias perspectivas (por exemplo, classificar “praia” pode ser a partir de uma perspectiva global, pessoal, microscópica etc.);
- Falta de transitoriedade. Ou seja, permite classificar uma entidade apenas numa única categoria, no entanto, nem sempre os atributos de uma entidade são comuns a todas as pessoas (por exemplo, se A é uma subclasse de B, e B é uma subclasse de C, então A também é uma subclasse de C. No entanto, a percepção humana em relação aos fenómenos não funciona assim);
- Regras restritas para inclusão de classes. As entidades nem sempre apresentam todos os critérios necessários e suficientes, como também podem ser melhor representadas por uma classe do que por outra. Entidades numa classe podem partilhar atributos em comum, mas talvez nem todas partilham dos mesmos atributos (por exemplo, os pinguins não são exclusivamente membros da classe das aves. Eles podem, de facto, possuir todos os atributos necessários e suficientes para pertencer a esta classe, mas não de forma tão evidente quanto as características apresentadas por uma águia.).

## Árvores

Da mesma forma que em Hierarquias, a estrutura Árvores divide e subdivide as classes baseada em regras específicas, onde cada objecto de classificação é elemento de uma classe. No entanto, ela não assume as regras de hereditariedade, ou seja, as classes são definidas por características específicas e mutuamente disjuntivas (Kwasnik 1999), (Fettke and Loos

2003). As suas principais utilidades são: a) apresentar informações sobre as entidades e os seus relacionamentos ao longo de uma dimensão de interesse; b) evidenciar o posicionamento relativo entre as entidades; c) demonstrar a frequência das entidades em qualquer nó. Por outro lado, são limitadas naquilo que podem representar, principalmente em termos de conhecimento sobre entidades numa mesma classe (Kwasnik 1999).

## **FIGURA 2 – Modelo de estrutura em árvore**

---

INESC-Porto

Unidade de Engenharia e Sistemas de Produção (UESP)  
    Redes de Cooperação Empresarial  
    +  
Unidade de Inovação e Transferência de Tecnologia (UITT)  
    Processos de Gestão da IDI  
    +  
Unidade de Optoelectrónica e Sistemas Electrónicos (UOSE)  
    Instrumentação  
    Telecomunicações  
    +  
Unidade de Sistemas de Energia (USE)  
    +  
    +  
Unidade de Sistemas de informação e Comunicação (USIC)  
    +  
    +  
Unidade de Telecomunicações e Multimédia (UTM)  
    Serviços Audiovisuais  
    +

---

No caso da estrutura em árvore, ao contrário de hierarquias, as subclasses não são “*um tipo de*” das super-classes, mas “*uma parte de*”. No exemplo do INESC-Porto, o elemento “Redes de Cooperação Empresarial” é uma parte de “Unidade de Engenharia e Sistemas de Produção”, e não “*um tipo de*”.

Vantagens:

- Informações completas e abrangentes;
- Regras de distinções sistemáticas e previsíveis. A estrutura geral é determinada pelos relacionamentos entre as entidades;

- Ordem de citação. É importante para decidir a ordem na qual as regras de distinção serão citadas (por exemplo, “vertebrados” ou “invertebrados”);
- Destaca os relacionamentos de interesse. Define a importância e a hierarquia entre as entidades;
- Revela a distância entre as entidades;
- Relativa frequência de entidades. Quando diversas entidades se agrupam sob um determinado rótulo de classificação, torna-se possível a criação ou descoberta de novas regras para as distinguir. Quando o conjunto de entidades é pequeno, elas tendem a ser tratadas como se fossem todas iguais.

#### Desvantagens:

- Rigidez. Não permite a adição de entidades fora do esquema em árvore previamente definido. Quando isto for necessário, a estrutura de entrada deverá ser repensada ou, até mesmo refeita;
- Fluxo de informação na vertical em ambos os sentidos;
- Perspectiva selectiva. Ao dar ênfase a determinados relacionamentos, pode mascarar ou não revelar outras relações igualmente interessantes.

#### Paradigmas

Cada objecto de classificação é representado por diversas características, as quais não precisam ser mutuamente disjuntivas (Fettke and Loos 2003). As entidades são descritas pela intersecção de dois atributos de uma só vez. A matriz resultante revela a presença ou ausência e a natureza de uma entidade na intersecção (ex.: relacionamento entre sexo [masculino, feminino] e parentesco [pais, irmãos, tios, primos]). Uma mãe, por exemplo, faz parte da categoria "feminino" e "pais". Esta estrutura é uma boa ferramenta para descobertas. Revela a presença ou a ausência de nomes de entidades definidas por pares de atributos. Ela pode ser utilizada para comparações e para a exposição de padrões e anomalias no que diz respeito à variedade e distribuição dos termos (Kwasnik 1999).

**FIGURA 3 – Modelo de estrutura paradigmas**

	Pais		Irmãos		Tios		Primos	
	Port.	Ing.	Port.	Ing.	Port.	Ing.	Port.	Ing.
Masculino	Pai	Father	Irmão	Brother	Tio	Uncle	Primo	Cousin
Feminino	Mãe	Mother	Irmã	Sister	Tia	Aunt	Prima	Cousin

Fonte: exemplo adaptado de (Kwasnik 1999)

Vantagens:

- Nomenclatura. Permite o estudo de padrões de nomenclatura. É frequentemente utilizada em estudos de terminologias;
- Distinção e falta de distinção. Permite identificar qual é o ponto em que o critério de intersecção possui termos distintos (por exemplo, no inglês não há distinção de sexo para o termo *cousin*, enquanto no português ele é descrito como primo (no masculino) e prima (no feminino));
- Padrões de similaridade e diferença. Em termos de criação de conhecimento, a paradigmas frequentemente fornece uma ferramenta heurística para a descoberta de regularidades em padrões de distinção. Embora cada cultura tenha uma grande quantidade de sobreposições em terminologias equivalentes, existem algumas diferenças subtis que podem ter explicações históricas ou outras;
- Células vazias. Permite investigar as razões para a falta de um termo nas células de atributos.

Desvantagens:

- Requer o conhecimento do domínio. A fiabilidade de representação desta estrutura pode ser comprometida caso as dimensões sejam triviais e não reflectam os conceitos fundamentais;
- Limite de perspectivas. Um conjunto de dimensões bem escolhido pode produzir uma descrição válida, como também um filtro que limita o âmbito do que pode ser visualizado (ex.: uma forma de definir os relacionamentos de parentesco é através de ligações de sangue ou casamento, sendo o sexo (masculino, feminino) como dimensão de distinção. No entanto, esta perspectiva não aborda afinidades por adopções legais e outras formas socialmente definidas);

- Poder ilustrativo limitado. Pelo facto de invocar dimensões somente em pares, raramente descreve completamente um fenómeno.

### Análise Facetada

A análise facetada não é uma estrutura de representação, mas uma abordagem para o processo de classificação. Nela, cada objecto é classificado de acordo com diferentes pontos de vista, denominados facetas (Fettke and Loos 2003). Garantir classificações flexíveis, ou seja, adaptáveis a novos fenómenos é o seu principal desafio. Para Ranganathan *apud* (Fettke and Loos 2003), precursor desta abordagem, qualquer entidade complexa poderia ser vista a partir de determinado número de perspectivas, ou facetas.

Actualmente, a análise facetada tem sido reinterpretada e utilizada em diversos campos, tal como em softwares para computadores, patentes, livros e objectos de artes (ex.: tesouros de artes e arquitectura). Ela permite algumas maneiras sistemáticas de visualização do mundo sem a necessidade de uma estrutura interna madura e estável (Kwasnik 1999).

**FIGURA 4 – Modelo da estrutura análise facetada**

Item	Autor	Ano	Assunto	Tipo de recurso	Idioma
1	Almeida Bax	2003	Ontologia	Artigo científico	Português
2	Al-Khalifa	2007	Taxonomia Classificação	Tese	Inglês
3	Caplan	2003	Metadados	Livro	Inglês
...	+	+	+	+	+

Um exemplo simples é apresentado na figura acima. Neste caso, os recursos de informação podem ser classificados até cinco facetas (autor, ano, assunto, tipo de recurso e idioma), onde cada uma pode receber a quantidade de atributos que for necessária. O item 2, por exemplo, aparecerá nos resultados de uma pesquisa feita tanto pelo assunto “taxonomia” quanto pelo assunto “classificação”. No entanto, não aparecerá nos resultados caso seja incluído mais um filtro pelo idioma “português”.

Na análise facetada, cada faceta pode ser desenvolvida e estruturada de forma independente. Por exemplo, a faceta “Ano” pode ser desenvolvida como uma linha do tempo e, a faceta “Assunto” pode ser desenvolvida de uma forma hierárquica.

#### Vantagens:

- Não requer conhecimento completo. Não é preciso conhecer toda a extensão ou relacionamentos da entidade para acomodá-la no esquema;
- Hospitalidade. Pode facilmente acomodar novas entidades;
- Flexibilidade. Uma vez que descreve cada objecto por um conjunto de atributos independentes, estes atributos podem ser infinitamente invocados e combinados de maneira flexível;
- Expressividade. Cada faceta está livre para incorporar o vocabulário e a estrutura que melhor atenda à representação do conhecimento;
- Não requer teoria sólida. Pelo facto de não ter uma estrutura global, não precisa de nenhuma “cola teórica” para mantê-la em conjunto e para conduzir as regras de associação e distinção;
- Pode acomodar uma variedade de estruturas e modelos teóricos. Por exemplo, ao analisar um excerto literário, uma faceta pode reflectir um modelo particular de géneros, outra um modelo de linguagens e assim por diante;
- Múltiplas perspectivas. Permite que uma entidade seja vista a partir de várias perspectivas.

#### Desvantagens:

- Dificuldade em estabelecer facetas apropriadas. Por não ser necessário conhecer o domínio nem os possíveis utilizadores, estabelecer as categorias fundamentais torna-se uma tarefa difícil;
- Falta de relacionamentos entre as facetas. Cada uma funciona como uma estrutura à parte;
- Díficil visualização. A partir de uma perspectiva visual, não pode ser exibida de tal forma que as suas entidades e relacionamentos sejam evidentes. Apenas pode ser vista ao longo de uma ou duas dimensões ao mesmo tempo.

Segundo (Su 2005), as quatro estruturas apresentadas por (Kwasnik 1999) podem ser agrupadas em apenas dois tipos. Para o autor, a estrutura hierarquia é um tipo especial de árvore e a estrutura paradigma é um tipo especial de análise facetada.



A principal diferença na estrutura da árvore e da análise facetada está no número de perspectivas possíveis para dividir o conhecimento no esquema de classificação. Enquanto na árvore apenas uma perspectiva é permitida, na análise facetada várias perspectivas podem ser incluídas (Su 2005).

### **2.1.2 Tipologia dos Esquemas de Classificação**

Os diferentes tipos de esquemas de classificação podem variar em âmbito, metodologia e características. Existem alguns tipos teoricamente possíveis, no entanto, uma classificação não está restrita a apenas um deles e pode ser ajustada a mais de uma categoria.

Estes tipos são (Koch, Brümmer et al. 1997):

- Por cobertura de assunto: geral ou específico;
- Por idioma: multilíngue ou monolíngue;
- Por geografia: global ou nacional;
- Por criação/suporte da organização: representante de uma organização legal existente ou desenvolvido por uma organização local para uso próprio;
- Por ambiente do utilizador: bibliotecas com suportes de publicações ou serviços de documentação disseminam pequenos resumos dos documentos (ex.: *abstracts e index databases*);
- Por estrutura: enumerativo ou facetado;
- Por metodologia: desenvolvido de acordo com uma estrutura geral do conhecimento e disciplinas científicas ou utilizando documentos classificados existentes (Koch, Brümmer et al. 1997); (Tristão, Fachin et al. 2004).

Os tipos de esquemas de classificação frequentemente mais utilizados são: a) universal; b) nacional; c) assunto específico; d) sistemas personalizados; e) adoção local de todos os tipos (Koch, Brümmer et al. 1997).

### **2.1.2.1 Esquema de Classificação Universal**

Com o rápido crescimento do conhecimento e o aumento do número de livros impressos, os Esquemas de Classificação Universais surgiram, no final do século XIX, para colmatar problemas de organização das bibliotecas (Koch, Brümmer et al. 1997). O termo “universal” é utilizado para esquemas que visam a inclusão de todos os assuntos, são geograficamente globais e possuem um âmbito multilingue (Uddin, Mezbah-ul-Islam et al. 2006). Os mais utilizados são o *Dewey Decimal Classification* (DDC), o *Universal Decimal Classification* (UDC) e o esquema de classificação da *Library of Congress* (LCC) (Koch, Brümmer et al. 1997).

Vantagens:

- Possível cobertura de todas as áreas temáticas;
- São amplamente apoiados: existe um interesse global em apoiar, desenvolver e manter;
- Podem ser mais familiares para os utilizadores do que os outros tipos de classificação: utilizadores regulares de bibliotecas podem estar mais familiarizados com estas classificações;
- Tem potencial para permitir acesso multilingue a uma colecção;
- Actualmente a maioria deles está disponível em formato digital.

Desvantagens:

- Falsa ontologia: imprimem uma preocupação geral em relação à imposição de uma falsa ordem no conhecimento;
- Custo na assimilação de novas áreas: tem dificuldade em reagir rapidamente ao aparecimento de novas áreas de estudo.

### **2.1.2.2 Esquema de Classificação Nacional**

São sistemas universais na cobertura de assuntos, mas normalmente destinados à utilização num único país. Alguns exemplos no contexto europeu são o *Nederlandse Basisclassificatie* (BC) e o *Sveriges Allmänna Biblioteksforening* (SAB). A maioria das vantagens e desvantagens dos esquemas de classificação universais aplicam-se também aqui. No entanto, eles pos-

suem algumas características adicionais que talvez não os tornem a melhor opção para um serviço de *Internet* (Koch, Brümmer et al. 1997).

Vantagens:

- No país de origem, pode ser mais conhecido do que os esquemas universais.

Desvantagens:

- Em geral não são bem conhecidos fora do seu país de origem;
- O suporte fora daquele grupo de utilizadores nacional é limitado;
- Capacidade multilingue não é a sua principal preocupação, excepto nos países onde se fala mais do que uma língua;
- Susceptíveis a uma tendência geográfica. Num contexto internacional, isto poderia ser uma séria desvantagem.

### **2.1.2.3 Esquema de Classificação por Assunto Específico**

São esquemas desenvolvidos para atender a um determinado grupo de utilizadores. Tipicamente são utilizados com serviços de indexação e resumos, colecções especiais ou importantes periódicos e bibliografias numa disciplina científica. O *Iconclass* para as artes, o *National Library of Medicine* (NLM) para a medicina e o *Engineering Information* (Ei) para a engenharia, são alguns exemplos. No caso da classificação de recursos da *Web*, estes esquemas são indicados apenas quando existe uma estrutura bem estabelecida (Koch, Brümmer et al. 1997).

Vantagens:

- Permitem uma estrutura e uma terminologia muito mais próxima à disciplina e, comparado aos esquemas universais, podem ser mais rapidamente actualizados.

Desvantagens:

- Dificil cooperação entre os serviços de classificação das diferentes áreas temáticas;
- Para utilizadores de outras áreas temáticas, aprender a estrutura do esquema de classificação pode ser uma tarefa muito difícil;

- Podem incluir alguns tópicos marginais que não serão adequadamente contemplados pelo próprio esquema.

#### **2.1.2.4 Esquema de Classificação Personalizado**

Os esquemas de classificação personalizados são concebidos para serem utilizados num determinado serviço. Alguns sites da *Web*, como o Yahoo!, por exemplo, através de uma ontologia própria, tentam organizar o conhecimento existente na *Internet* através do seu próprio esquema de classificação (Koch, Brümmer et al. 1997).

Vantagens:

- São relativamente flexíveis e fáceis de alterar;
- Podem muito rapidamente absorver novas áreas de interesse.

Desvantagens:

- Aumentam os problemas de subjectividade na classificação e podem levar a uma falta de coerência;
- Provavelmente não serão tão bem conhecidos pelos utilizadores como os esquemas universais;
- Como é auto-elaborado, poderia necessitar de revisões frequentes, com poucas possibilidades de cooperação. O encargo económico recairia inteiramente no criador do esquema.

#### **2.1.3 Estruturas de Classificação utilizadas na Web**

Além das classificações para bibliotecas, os sistemas de classificação têm sido desenvolvidos e utilizados em diversos contextos. A maioria deles segue mais ou menos um mesmo tipo de estrutura, pois agrupam entidades baseadas nas suas propriedades, estruturam-nas em alguns níveis hierárquicos e utilizam símbolos para manter a hierarquia. Na *Web*, a maioria das estruturas de classificação utilizadas são hierárquica-enumerativa e analítica sintética (facetada) (Uddin, Mezbah-ul-Islam et al. 2006).

As classificações na *Internet* podem ser agrupadas em três principais estruturas, sendo a classificação hierárquica-enumerativa, classificação facetada e classificação orientada ao utilizador (Uddin, Mezbah-ul-Islam et al. 2006):

- **Classificação hierárquica-enumerativa:** estrutura hierárquica onde o conhecimento é dividido progressivamente em categorias pequenas e mais específicas, sendo que cada objecto está tipicamente situado apenas numa categoria. Este esquema *top-down* atribui nomes e enumera cada assunto, normalmente numa ordem sistemática;
- **Classificação facetada:** abordagem *bottom-up* onde cada objecto é identificado com um determinado conjunto de atributos e valores (designados facetas). A sua organização vai emergindo de acordo com a forma utilizada para aceder à informação. É um método de descrição multidimensional onde os recursos de informação são organizados por assuntos, atributos ou características comuns. Na *Web* pode ser uma boa ferramenta para a recuperação de informação, pois fornece uma estrutura do conhecimento mais legível e compreensível. Isto auxilia no desenvolvimento de uma arquitectura da informação mais próxima das necessidades dos utilizadores da *Internet*;
- **Classificação orientada ao utilizador:** classificação informal através da prática de categorização colaborativa de recursos da *Web*. Para armazenar e recuperar conteúdos da *Internet*, os utilizadores escolhem e atribuem livremente os seus próprios descritores (conhecidos como *tags*). Esta classificação, conhecida como folksonomia, não possui hierarquia entre os descritores atribuídos, como também não especifica directamente nenhuma relação entre eles. A partir do agrupamento de termos baseados em URLs comuns, pode definir-se o relacionamento entre as *tags*. A folksonomia produz resultados que reflectem mais fielmente a população do modelo conceptual da informação.

Num ambiente *Web*, os esquemas de classificação podem oferecer uma solução para melhorar o acesso às informações. A sua utilização em sítios da *Internet* pode trazer alguns benefícios, tais como (Koch, Brümmer et al. 1997):

- Navegação: as listas de assuntos classificados são facilmente navegáveis num ambiente *online*;
- Ampliação (favorável a recuperação) ou restrição (favorável a precisão) das pesquisas;

- Contexto: esquemas de classificação contextualizam os termos utilizados na pesquisa;
- Permitem o acesso multilingue a uma colecção;
- Partição e manipulação de uma base de dados: grandes listas de classificação podem ser divididas logicamente em pequenas partes;
- Possibilita melhorar a navegação e pesquisa de assuntos em bases de dados;
- Mesmo que o esquema de classificação seja actualizado, não é necessário reclassificar todos os recursos;
- Podem facilmente tornar-se familiares;
- A maioria está disponível em formato digital.

No entanto, podem apresentar igualmente algumas desvantagens (Koch, Brümmer et al. 1997):

- Colecções de materiais relacionados podem ser divididas. Isto pode ser parcialmente superado através de referências cruzadas;
- Subdivisão de classes incoerente pode dificultar a navegação;
- Dificil assimilação de novas áreas de interesse: uma vez que são normalmente actualizadas através de processos formais feitos por entidades organizadas, frequentemente revelam dificuldades em reagir a novas áreas de estudo;
- Custo e complexidade: a utilização e a manutenção podem implicar grandes custos. Além disso, a sofisticação do esquema de classificação pode produzir complexidade entre o utilizador e o sistema de informação (Uddin, Mezbah-ul-Islam et al. 2006).

Em relação à utilização de estruturas de classificação na arquitectura de informação da *Web*, (Uddin, Mezbah-ul-Islam et al. 2006) relatam que o seu correcto uso pode proporcionar uma área de trabalho amigável ao utilizador, servir como um método de recuperação de informação mais económico e também assegurar uma óptima utilização da informação corporativa.

Em serviços da *Internet*, os esquemas de classificação são vantajosos quando ajudam na navegação, possibilitam pesquisas mais abrangentes ou específicas, contextualizam os termos

utilizados na busca, permitem acesso multilingue às colecções de materiais, partição e manipulação de uma grande base de dados (Koch, Brümmer et al. 1997).

## **2.2 Alguns Sistemas de Organização do Conhecimento**

Existem vários sistemas e ferramentas de apoio à classificação e organização do conhecimento. Do ponto de vista dos utilizadores, alguns podem ser relativamente simples, no entanto, do ponto de vista dos seus organizadores, podem representar uma difícil tarefa. Entre estas ferramentas podemos enumerar algumas: vocabulários controlados, cabeçalhos de assuntos e tesouros.

### **2.2.1 Vocabulário Controlado**

Um vocabulário controlado é constituído por termos, onde um termo é um nome particular para um determinado conceito. Normalmente não existem termos duplicados, referindo-se apenas a um simples assunto. Ele não é constituído directamente por conceitos, mas por um conjunto de termos indexados, ou, assuntos utilizados para a classificação (Garshol 2004).

É uma lista estabelecida de termos preferidos que um catalogador ou indexador deve seleccionar enquanto atribui assuntos ou descritores para um registo bibliográfico, com o objectivo de indexar o conteúdo dos trabalhos existentes num catálogo de biblioteca, índice ou base de dados bibliográfica. Sinónimos são incluídos como vocabulários de ligação, com as instruções *see* (direcciona para o termo correcto) ou *use* (indica que é o cabeçalho válido). O vocabulário controlado é normalmente listado alfabeticamente numa lista de cabeçalho de assuntos ou tesouro de termos indexados (Reitz 2007).

### **2.2.2 Cabeçalho de Assuntos**

O cabeçalho de assuntos é constituído por um conjunto de termos que representam os itens numa colecção. Pese embora a possibilidade de se tornar extenso e cobrir um vasto leque de assuntos, a sua estrutura é geralmente muito superficial e possui uma hierarquia limitada. A tendência é que eles sejam coordenados, com regras de ligação entre os termos para que estes forneçam os conceitos mais específicos para cada assunto. O *Medical Subject Headings (MeSH)* e o *Library of Congress Subject Headings (LCSH)* são exemplos de cabeçalhos de assuntos utilizados na *Internet* (Hodge 2000).

Para auxiliar na identificação dos temas, um cabeçalho de assuntos pode ser dividido em sub-classes, como em **Computadores – Internet – Directórios**, ou pode ser incluído um qualificador entre parênteses para clarificação semântica, como em Apple (Informática) (Reitz 2007).

### **2.2.3 Tesouros**

Os tesouros – um tipo de cabeçalho de assuntos mais avançado – são baseados em conceitos e apresentam relacionamentos entre os termos. Normalmente, estes relacionamentos incluem hierarquia, equivalência (sinónimos) e associação ou parentesco. Os relacionamentos geralmente são representados pelas notações BT (*broad term* [termo geral]), NT (*narrow term* [termo específico]), SY (*synonym* [sinónimo]) e RT (*associative or related term* [termo associado ou relacionado]). Dois exemplos de tesouros *online* são o *Aquatic Sciences and Fisheries Thesaurus* e o *National Aeronautic and Space Administration (NASA) Thesaurus* (Hodge 2000).

As vantagens obtidas por esta abordagem são a possibilidade de pesquisa directa de informações catalogadas manualmente neste cabeçalho de assunto, e a capacidade de navegação numa esfera de tópicos relevantes através das relações entre termos ligados, porém, diferentes (Beck and Pinto 2002).

## **2.3 Ontologias**

O termo ontologia tem origem no grego "*ontos*" (ser) e "*logos*" (palavra) (Sowa 2005). Ele foi emprestado da Filosofia, onde a ontologia é uma descrição sistemática da Existência. Ela é uma especificação explícita de uma conceptualização. Para sistemas baseados no conhecimento, o que "existe" é exactamente o que pode ser representado (Gruber 1993). É um catálogo de tipos de coisas que parte do princípio que estas existem ou existirão num domínio de interesse (Sowa 2005).

Sob a perspectiva da Inteligência Artificial e Representação do Conhecimento, (Gruber 1993) p.2) relata que:

"[...] we can describe the ontology of a program by defining a set of representational terms. In such an ontology, definitions associate the names of entities in the universe of discourse (e.g., classes, relations, functions, or other objects) with human-readable text describing



what the names are meant to denote, and formal axioms that constrain the interpretation and well-formed use of these terms."

As ontologias têm sido apresentadas para resolver os problemas que surgem da utilização de diferentes terminologias para referenciar o mesmo conceito, ou da utilização do mesmo termo para referenciar diferentes conceitos. Elas especificam um vocabulário comum entre diferentes sistemas, ou seja, tendem a representar o conhecimento que é mais ou menos consensual numa comunidade de pessoas. Informalmente, uma ontologia consiste num conjunto de termos e num conjunto de condições impostas de maneira a que estes termos possam ser combinados (Beck and Pinto 2002). Ela estabelece critérios para distinguir vários tipos de objectos (concreto e abstracto, existente e não existente, real e irreal, dependente e independente) e os seus relacionamentos (relações, dependências e predicções) (Corazzon 2008).

As Ontologias podem ser classificadas em quatro tipos, sendo ontologias de representação, ontologias de alto nível, ontologias de domínio e ontologias de aplicação (Beck and Pinto 2002):

- **Ontologias de representação:** captura as primitivas representações utilizadas para formalizar o conhecimento numa dada família ou sistema de representação do conhecimento (ex.: definição de conceitos como classes, relações, funções etc.);
- **Ontologias de alto nível:** classifica as diferentes categorias de entidades existentes no mundo. Representa noções muito gerais que são independentes de um problema ou domínio. O conhecimento é aplicável através de domínios e inclui vocabulário relacionado às coisas, eventos, tempo, espaço e meteorologia;
- **Ontologias de domínio:** são específicas para um particular domínio. Descrevem um vocabulário relacionado com um domínio genérico, como aeronaves, elementos químicos etc.
- **Ontologias de aplicação:** conectam ontologias com métodos de resolução de problemas num sistema particular baseado no conhecimento.

## ***2.4 Taxonomias***

Numa definição geral, a taxonomia é a ciência da classificação. Mais especificamente, é a classificação de organismos vivos e extintos. O termo é derivado do grego *taxis* ("organiza-

ção”) e *nomos* (“ciência”). É a metodologia e o princípio de organizar os tipos de plantas e animais numa hierarquia de grupos superiores e subordinados (Britannica 2008).

Além da biologia, outras áreas de estudo passaram a utilizar a taxonomia para organizar o conhecimento, como por exemplo a informática e tecnologias da informação, ciências muitas vezes aplicadas às organizações. Alan Gilchrist (Gilchrist 2003) aponta alguns factores que possivelmente influenciaram estas áreas de interesse:

- **Excesso de informação:** motores de pesquisa convencionais são considerados insuficientes para grandes bases de dados, tornando evidente a necessidade de filtros e ferramentas de ajuda para complementar as pesquisas dos utilizadores;
- **Conhecimento da informação:** a maioria dos utilizadores não sabem como pesquisar a informação, levando ao desperdício de tempo e à perda de informações úteis;
- **Terminologia organizacional:** a maioria das informações consumidas pelas organizações são criadas internamente, e os tesouros e as classificações publicadas nem sempre reflectem esta linguagem particular;
- **Falta de estrutura das organizações:** fusões e aquisições entre organizações, parcerias através de extranets e o estabelecimento de comunidades virtuais, têm criado problemas culturais, principalmente na etapa de implementação.

Segundo (Kremer, Kolbe et al. 2005), as taxonomias servem de suporte à pesquisa, navegação, armazenamento e serviços de comunicação entre pessoas e sistemas. Elas são o resultado da gestão terminológica, que é a soma de unidades organizacionais, processos e instrumentos que apoiam a criação e gestão de termos e classificações numa área específica.

Com base no artigo de (Gilchrist 2001), que contém um estudo sobre o uso da taxonomia por grandes organizações, Alan Gilchrist (Gilchrist 2003) identificou, pelo menos, cinco diferentes significados para a palavra taxonomia:

- **Directórios Web:** são uma forma de classificação normalmente utilizada na *Internet*, e de maneira crescente, embora com menor frequência, em intranets. É um modo de organização constituída por assuntos e sub-assuntos, onde cada nível não precisa de estar necessariamente numa hierarquia com algum sentido convencional. Os termos podem repetir-se em diferentes níveis, permitindo alternativas para o pesquisador;

- **Taxonomias para suporte de indexação automática:** possui um algoritmo que utiliza regras de extracção automática de termos e palavras-chave que podem ou não estar presentes nos documentos. Esta abordagem é atractiva, onde é grande o fluxo de informações e a indexação manual é economicamente impossível;
- **Taxonomias criadas por categorização automática:** são sistemas capazes de analisar textos, criar categorias automaticamente e analisar documentos de acordo com as categorias criadas. As taxonomias podem ser apresentadas como os directórios *Web* ou em mapas de duas dimensões, onde os termos relacionados estão ligados ao termo seleccionado que aparece no meio do mapa;
- **Formulário de entrada e filtros:** tipo de taxonomia utilizada em formulários de pesquisa. O utilizador pode navegar na hierarquia em qualquer sentido, bem como saltar para outros termos relacionados. O termo seleccionado na pesquisa é enviado à Internet ou intranet. Esta taxonomia é um tesouro formatado para permitir uma fácil navegação;
- **Taxonomias corporativas:** usualmente é realizada de forma manual e resulta num mapa de alto nível dentro do portal da empresa, orientando os colaboradores para o capital intelectual da organização, que pode abranger informações sobre si própria, tais como: processos de negócios, métodos, guias, normas, informações sobre colaboradores e pessoas associadas à organização, bem como conteúdos contidos em vários repositórios.

As taxonomias podem utilizar uma combinação de classificação e técnicas de tesouros aplicadas a um vasto leque de tipos de objectos.

### 3 FOLKSONOMIA: SISTEMA INFORMAL DE ORGANIZAÇÃO DO CONHECIMENTO

É secular a utilização da classificação para contextualizar e direccionar qualquer aspecto do conhecimento humano. Especialistas, como bibliotecários e cientistas da informação, têm feito um grande esforço para criar poderosos e apropriados sistemas de classificação (Quintarelli 2005). Com a rápida explosão de informação que prolifera na *Web*, criar um sistema profissional de metadados para esses recursos exige muito tempo, dinheiro e esforço. Uma forma de classificação, comum na organização de conteúdos electrónicos, é atribuir termos que os identifiquem e ajudem numa futura navegação.

Embora não seja uma ideia inovadora, devido aos actuais serviços da *Web*, que utilizam um mecanismo intuitivo para descrever os seus recursos, uma forma colaborativa desse processo está cada vez mais popular na Internet, a qual muitos chamam de folksonomia (Al-Khalifa 2007); (Vanderlei 2006).

São muitos os sinónimos que podemos encontrar para o termo folksonomia, tais como social tagging, collaborative tagging, social classification, social indexing, folk classification, ethno-classification, distributed classification, open tagging, free tagging, distributed classification, mob indexing, federated tagging, lazy tagging, tagsonomy, tagonomy, post coordinate indexing, collective indexing, user-generate tagging entre outros (Wikipedia 2007); (Hammond, Hannay et al. 2005); (Al-Khalifa 2007).

O termo folksonomia, o mais conhecido e utilizado, levanta algumas discussões pois alguns autores discordam do uso da palavra “taxonomia” incutida no termo “folksonomia”. Para (Merholz 2004), por exemplo, a taxonomia é mais precisa e possui uma hierarquia (na maioria das vezes imposta), o que não ocorre num sistema de *tagging*, como a folksonomia. Segundo o autor, o termo mais apropriado seria "*ethnoclassification*".

O termo *folksonomy* surge em Julho de 2004, no seguimento de uma pergunta colocada na lista de discussão da IA Institute (então designado Asylomar Institute for Information Architecture – AIFIA). Gene Smith, autor do *post*, questionava o tipo de classificação de informação de serviços como o Furl, Flickr e Delicio.us. Várias seriam as respostas encontradas a esta questão, e seria inspirado na resposta de Eric Scheid, da Ironclad Information Architecture, que escrevera "*folk classification*", que Thomas Vander Wal concluiu "*So the user-created bottom-up categorical structure development with an emergent thesaurus would become a Folksonomy?*". Em de Agosto de 2004, Gene Smith mencionaria no seu blog "*Folksonomy: Social Classification*", o qual atiraria a expressão folksonomia para fora da lista de discussão do IA (Wal 2007).

A explicação etimológica da palavra “folksonomia” configura a união das palavras *folks* (de origem germânica que significa povo, gente) e taxonomia (do Grego verbo *tassein* = "para classificar" e *nomos* = "lei, ciência, administrar").

A folksonomia tem pouco a ver com a taxonomia, pois esta última refere-se a uma forma de classificação hierarquizada, enquanto a folksonomia estabelece categorias (cada *tag* é uma categoria) que teoricamente são iguais umas às outras. A parte "*-nomy*", comum nas duas palavras, vem do grego "*nomia*" e significa "gestão".

Assim sendo, folksonomia poderia ser traduzida superficialmente como "povos gestão". (Wikipedia 2007).

A folksonomia é o resultado da livre criação de *tags* pessoais para informações e objectos (algo que possua uma URL) para a sua própria recuperação. Este processo de *tagging* é feito num ambiente social, normalmente distribuído e aberto por outros utilizadores. A folksonomia é criada a partir do acto de incluir *tags* pela pessoa que consulta a informação (Wal 2007). Ela é um

"[...] sistema democrático de *tagging*, que reflete a opinião do público em geral sobre determinado objeto, associado à identificação das pessoas que fazem essa classificação, introduzindo uma abordagem distribuída, inovadora, baseada em uma classificação social." (Vanderlei 2006).

Para (Quintarelli 2005), as folksonomias não são teorias ou estratégias *top-down*, mas nasceram de um recurso introduzido por *softwares* como Del.icio.us e Flickr, e de utilizadores destas plataformas para criarem *tags* para os seus conteúdos. A folksonomia requer pessoas para associarem as suas próprias palavras-chave aos conteúdos, e para posterior recuperação. O seu poder está no acto de agregação e não simplesmente na criação de *tags*. Não são apenas pessoas a criar *tags* para uso pessoal, mas também uma agregação das informações que elas fornecem.

(Damme 2006) p.31) define a folksonomia como

"[...] a social classification mechanism created by users which results in a bottom-up taxonomy. In this mechanism everyone can freely choose keywords for categorizing or describing their content. However, it is more than simply labeling or tagging objects or information, because this labeling process occurs in a social environment: everybody has access to the tagged content and the used tags and tag feedback, based on the input of others users, is generated."

Constata-se a utilização pelos autores da palavra "*tag*" nas suas definições. Uma *tag* pode ser descrita como "[...] keywords, category names, or metadata. In essence, a tag is simply a freely chosen set of textual keywords." (Guy and Tonkin 2006). Para os autores, como ela não é criada por especialistas e em nenhum momento segue orientações formais, pode ser categorizada com qualquer palavra que define um relacionamento entre o recurso online e um conceito na mente do utilizador. Assim, com a combinação de adjectivos sociais e colaborativos, um processo de *tagging* pode ser sinónimo para a folksonomia (Damme 2006).

### **3.1 Características das Folksonomias**

Tradicionalmente, a classificação ou indexação de conteúdos é feita por um ou vários especialistas (como no caso dos repositórios, que permitem que componentes da sua colecção sejam classificados por palavras-chave), ou deriva do material fornecido pelos autores dos componentes. Para um bom uso do sistema, este processo exige que se tenha uma vasta visão comum do domínio, e utilizadores que entendam as categorias e a estrutura de classificação. Pelo contrário, a folksonomia permite que qualquer pessoa atribua livremente palavras-chave ou *tags* aos objectos, de forma não hierárquica e sem relação entre os termos no vocabulário (Vanderlei 2006).

Um dos aspectos mais importantes da folksonomia está em permitir que os próprios utilizadores atribuam descritores às informações e aos objectos para a sua própria recuperação. Cada pessoa tem total autoridade para escolher e atribuir as suas próprias *tags*. Neste processo de classificação, existem três elementos distintos de dados: a *tag*, o objecto a ser marcado e o indivíduo que está a fazer a marcação. Entretanto, ao permitir a identificação dos autores das *tags* e discerni-las em diferentes informações e objectos, a folksonomia, além um processo de classificação, torna-se também um processo social, pois possibilita que as pessoas encontrem tanto indivíduos como comunidades com o mesmo interesse (Wal 2006).

A *tag cloud* (nuvem de *tags*) é outra importante funcionalidade da folksonomia. Ela é um conjunto de *tags* relacionadas que possuem pesos correspondentes. Tipicamente tem entre 30 e 150 *tags*. Os pesos são representados pelo tamanho das fontes e outras características visuais. Normalmente, as *tags* estão numa ordem alfabética e as mais utilizadas são representadas por fontes largas ou destacadas. Uma *tag clouds* normalmente é interactiva: as *tags* são *atalhos* que levam o utilizador a uma colecção de itens que estão associados com aquela *tag* (Wikipedia 2007); (Vanderlei 2006).

FIGURA 5 – Representação de uma tag cloud



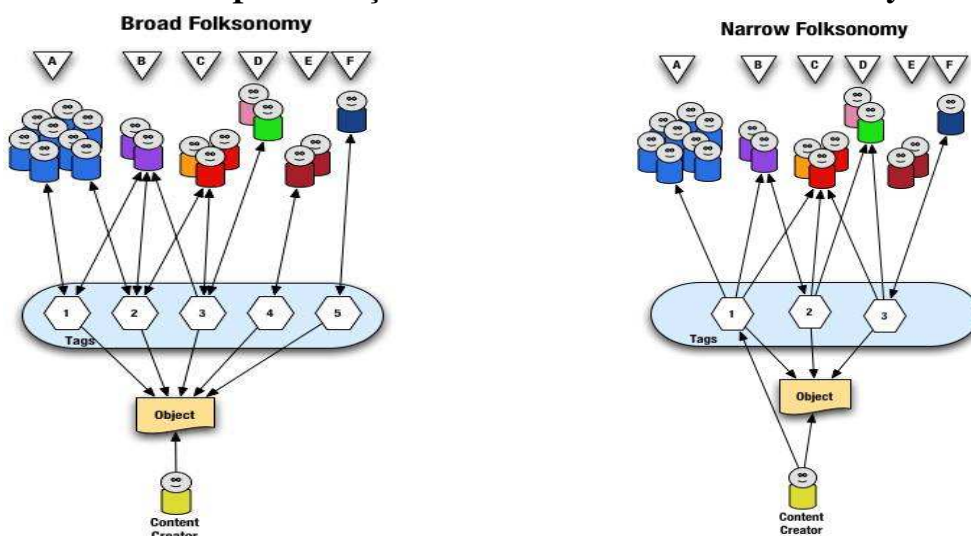
Fonte: (CREMONI 2008)

Um factor importante, talvez um dos grandes impulsionadores para a popularidade da folksonomia, é que tudo é feito de maneira simples e com relativo baixo custo, principalmente se comparado ao esforço e ao custo de desenvolver e manter uma taxonomia.

### 3.2 Tipos de folksonomia

A folksonomia é, na maioria das vezes, também social. Isto permite aos utilizadores que partilham do mesmo vocabulário encontrar os mesmos objectos. É importante notar que as folksonomias trabalham melhor quando as *tags* utilizadas se inserem num vocabulário comum, do que através da terminologia de classificação atribuída por um único indivíduo. Esta classificação social pode ser dividida em dois tipos, sendo *broad folksonomy* (folksonomia abrangente) e *narrow folksonomy* (folksonomia restrita) (Wal 2005).

**FIGURA 6 – Representação da Broad e Narrow Folksonomy**



Fonte: (Wal 2005)

Na folksonomia abrangente várias pessoas podem atribuir as suas próprias *tags* e o seu próprio vocabulário ao mesmo objecto (Wal 2005). Esta categoria, que pode ser exemplificada pelo site del.icio.us, é caracterizada pela amplitude do número de *tags* autorizadas a referenciar um item, e pelo facto dos utilizadores poderem incluir *tags* já emitidas para conteúdos existentes (Auray 2007). Segundo (Dye 2006),

Users of these sites can see what other tags have been created for certain content and use this to get a broad idea of associated terms and tags. They're often referred to as social classifications, since seeing what other users are thinking about is as much a part of the site as finding what you need.

A folksonomia abrangente proporciona uma maneira de analisar as tendências de como os objectos estão a ser categorizados. Da mesma forma que permite identificar os termos mais utilizados e a tendência de classificação da maioria das pessoas, possibilitando também a identificação de uma minoria que utiliza termos mais específicos para um determinado objecto. (Wal 2005).

Ao contrário da folksonomia abrangente, onde é permitida uma livre atribuição de descritores aos objectos, na folksonomia restrita apenas uma ou algumas pessoas podem adicioná-los. Frequentemente, as *tags* iniciais são atribuídas pelos próprios criadores dos objectos (Wal 2005). Este tipo de folksonomia é caracterizado por determinar um limite no número de *tags* a serem atribuídas, como ocorre, por exemplo, no site Flickr (Auray 2007). A sua maior valia está nos textos das *tags*, pois eles auxiliam na localização de objectos que não eram



detectáveis apenas com a ajuda de ferramentas de pesquisa, como no caso de imagens e fotografias (Wal 2005).

Para exemplificar esta mais valia, Jessica Dye (Dye 2006) coloca uma questão pertinente: “Even if Google could tell you that a particular picture contained a human, how would it know it was your Aunt Betty, or that it was taken in the summer of 2005?”

Resumidamente, enquanto que na folksonomia restrita somente o próprio criador do recurso pode atribuir-lhe *tags*, na folksonomia abrangente qualquer pessoa pode atribuir *tags* a qualquer coisa.

### **3.3 Aspectos Positivos e Negativos**

Muito tem sido discutido a respeito das vantagens e desvantagens da folksonomia (Mathes 2004); (Quintarelli 2005); (Kroski 2005); (Golder and Huberman 2006); (Hayman 2007); (Brouwers, Nawaz et al. 2007). Embora seja um fenómeno popular, ainda não são claros todos os seus potenciais benefícios. O facto é que um grande grupo de pessoas está a utilizá-la como uma alternativa à organização e recuperação de informações, que por sua vez, tem sido efectiva como um sistema de pesquisa de informação (Mathes 2004).

Autores como (Mathes 2004), (Quintarelli 2005) e (Kroski 2005) discutem sobre alguns aspectos positivos da folksonomia:

- Nem todas as limitações da folksonomia são defeitos, podendo ser vista como uma escolha de recursos;
- É simples, emergente e interactiva. A folksonomia tem uma grande capacidade de combinar as reais necessidades dos utilizadores com a linguagem, e não propriamente com a sua precisão;
- Oferece usabilidade, ou seja, não exige muita aprendizagem e é extremamente compreensível para o utilizador;
- Não é hierárquica e não possui nenhuma relação directamente especificada para os termos. Entretanto, gera automaticamente relações entre as *tags*, através do agrupamento baseado em URL's comuns;

- É serendipista<sup>1</sup>, ou seja, constitui um recurso extraordinário que, através de conjuntos de *tags* relacionadas, permite encontrar coisas ao acaso, que, sem este recurso, talvez nunca pudessem ser reveladas;
- É democrática. Permite que qualquer pessoa inclua algo num todo. Devido à sua natureza, estimula a personalização das *tags* para que estas auxiliem numa futura recuperação dos itens de interesse;
- Não é restritiva. Não há nenhum controlo terminológico centralizado, pois esta gestão é feita pelos próprios utilizadores. Desta forma, ideias de interesse de uma pequena fracção da população podem surgir e despertar a atenção de toda a massa. A folksonomia reflecte o vocabulário e as ideias dos utilizadores;
- Para os utilizadores do sistema, os custos globais de tempo, esforço e aspectos cognitivos são muito inferiores do que em esquemas hierárquicos de classificação e categorização complexos;
- Pode ser considerada uma ponte de baixo investimento entre a classificação pessoal e a classificação partilhada;
- É colaborativa. Cria um contexto de comunicação e partilha de informações. Mesmo sem a obrigatoriedade em partilhar, por ser um ambiente interactivo e por estar totalmente interligado, o acto do utilizador incluir *tags* a objectos não é apenas uma acção individual, mas também uma forma, às vezes inconsciente, de partilhar materiais com outros utilizadores;
- Diminui barreiras para a cooperação. Não é necessário um acordo entre grupos de utilizadores para definir uma hierarquia de *tags* ou taxonomias detalhadas. É preciso apenas que eles concordem no significado de uma *tag*, a qual é suficiente para rotular materiais semelhantes;
- Feedback imediato. Logo que uma *tag* é atribuída a um objecto, o grupo de itens que possuem a mesma *tag* é apresentado. Caso não seja o resultado esperado, a sua alteração é simples e intuitiva. O grande paradigma não está apenas em encontrar mas também em descobrir informações. Naturalmente, a folksonomia

---

<sup>1</sup> Serendipidade, também conhecido como Serendipismo, Serendiptismo ou ainda Serendipitia, é um neologismo que se refere às descobertas afortunadas feitas, aparentemente, por acaso Wikipedia. (2008). "Serendipidade." Retrieved February 10th, 2008, from <http://pt.wikipedia.org/wiki/Serendipismo..>

é menos precisa do que vocabulários controlados ou taxonomias, mas, por outro lado, incentiva a que os utilizadores naveguem, investiguem e descubram conteúdos de forma simples e agradável;

- Ao se tratar de eixos tradicionais de comparação, metadados bem concebidos são melhores do que uma folksonomia. No caso da *Web* o ambiente faz a diferença. Por conter um grande volume de informações e por ser um ambiente extremamente dinâmico, é necessário que haja uma classificação rápida e, conseqüentemente, menos profissional. Neste caso, a folksonomia é o melhor que se tem até agora. Ela é o limite entre as tradicionais estruturas de classificação e a não classificação;
- Sistemas baseados em *tagging* oferecem uma facilidade e actualização que não são possíveis em sistemas com vocabulário controlado. Utilizadores criam *tags* tão rápido quanto criam conteúdos. Esta flexibilidade permite uma resposta rápida às mudanças de terminologias e aos eventos mundiais;
- A folksonomia proporciona uma vista sobre o comportamento dos utilizadores. Ela permite observar como eles atribuem *tags* aos seus recursos. Além disto, possibilita identificar quais são os tipos de categorias que têm surgido. O desenvolvimento de uma futura taxonomia *top-down* poderia surgir a partir da observação da linguagem e do comportamento dos utilizadores.

Por outro lado, a folksonomia apresenta também aspectos negativos (Quintarelli 2005); (Mathes 2004); (Golder and Huberman 2006); (Hayman 2007); (Brouwers, Nawaz et al. 2007):

- Imprecisão face à complexidade da língua, em função do comportamento do utilizador:
  - **Polissemia:** A palavra polissemia significa ("poly") vários ("semy") sentidos. Por ser um sistema de livre adição de termos, os utilizadores podem, em diferentes situações, aplicar termos para documentos de diversas maneiras. Por exemplo, o termo "janela" pode indicar tanto um buraco na parede de um imóvel quanto o painel de vidro que está dentro dele. Isto diminui a performance nos resultados da busca, devolvendo, muitas vezes, itens potencialmente não aplicáveis;
  - **Sinónimos:** não há controle de sinónimos. Sinónimos e termos semelhantes representam um grande problema para sistemas de *tagging*, pois a

incoerência dos termos utilizados podem tornar difícil para o utilizador a identificação dos itens relevantes recuperados numa pesquisa. Por exemplo, termos como "mac", "macintosh" e "apple" (que também pode referir-se à fruta maçã) são usados para descrever materiais da Apple Macintosh;

- **Variações de níveis básicos:** alguns factores de níveis básicos, tais como plural, singular, excertos de textos e ortografia podem também influenciar nos resultados das pesquisas. As palavras "gato" e "gatos", por exemplo, representam a mesma coisa, mas num sistema de *tagging* elas são distintas. Neste caso, uma pesquisa para um dos termos não retorna as informações relacionadas com ambas as terminologias;
- **Sensibilidade:** algumas *tags* podem também ser sensíveis em relação ao formato da letra, retornando diferentes resultados quando utilizadas letras minúsculas ou maiúsculas (por exemplo, "português" vs. "Português").
- Propõe *tags* não hierárquicas, o que torna difícil a elaboração de estruturas complexas de *tags* para descrever os seus objectos ou objectos de outros *sites*. Um mesmo item pode ser descrito de maneiras diferentes, variando de um nível de abstracção geral ao mais específico. Por exemplo, "gato", "leopardo" e "animal" podem ser *tags* adicionadas para um mesmo item;
- Tem um quociente de localização muito baixo. São excelentes para navegação e pesquisas casuais, mas não para abordagens ou pesquisas segmentadas, pois este não é o seu propósito;
- Alguns sistemas permitem somente uma palavra para as *tags*, o que dificulta a definição de conceitos complexos. Uma alternativa seria criar termos compostos por algumas palavras, no entanto, isto tornaria as *tags* termos altamente complexos e imperceptíveis, como também influenciaria negativamente na precisão das pesquisas;
- Maior dificuldade em encontrar assuntos menos comuns e populares, uma vez que o processo de classificação depende dos utilizadores;
- Os sistemas de *tagging* estão mais susceptíveis aos utilizadores menos éticos que propagam *links*, neste caso, *tags* maliciosas com a finalidade de corromper a credibilidade e fiabilidade do sistema.

Uma folksonomia representa o que há de melhor e pior na organização de informação. Apesar de ser naturalmente descontrolada e caótica, imprecisa e ambígua, está a incentivar os utilizadores a organizar informações segundo os seus próprios critérios, estando fundamentalmente direccionada para as suas necessidades e vocabulários, envolvendo-os directamente na organização do sistema (Mathes 2004).

A folksonomia é benéfica tanto na construção de um vocabulário pessoal quanto em aspectos sociais, auxiliando os indivíduos a encontrar e agrupar uma gama de objectos na *Web* da sua esfera de interesses. Um factor importante a observar para quem inclui *tags* para um objecto é saber como elas serão pesquisadas. Isto porque um processo de *tagging* nem sempre é uma folksonomia (pois a folksonomia representa um processo distribuído e aberto a todos), mas apenas *tagging* (que é apenas o processo de atribuir um termo ou uma palavra-chave a um objecto). Criar *tags* para si próprio é útil, mas não se aproxima dos benefícios de tags abertas a todos, pois uma folksonomia pode permitir conexões através de culturas e disciplinas. (Wal 2005).

Se por um lado uma *tagging* colectiva, que é o caso da folksonomia, pode acentuar os problemas associados às fronteiras linguísticas e cognitivas, e produzir um sistema de classificação extenso, o qual consiste em categorias pessoais desestruturadas, por outro, possibilita uma aprendizagem mútua através da distribuição e organização de informações (Golder and Huberman 2006).

## 4 FOLKSONOMIA EMPRESARIAL

Hoje em dia, uma empresa diferencia-se da concorrência pelo seu conhecimento, e não pela sua produção. Isto quer dizer que as empresas que não investirem em criação e manutenção de conhecimento estão destinadas a desaparecer do mercado (Damme 2006).

Para Peter Drucker, em (Harris 1993) p.120),

"Knowledge is power, which is why people who had it in the past often tried to make a secret of it. In post-capitalism, power comes from transmitting information to make it productive, not from hiding it."

Actualmente, existem diversas ferramentas que permitem gerar e partilhar informações electrónicas de forma rápida e fácil. O resultado disso é um processo acumulativo que cresce exponencialmente a cada dia, criando, assim, um excesso de informação. O facto é que apesar de existirem várias informações valiosas nesse contexto, elas não estão estruturadas, o que definitivamente torna difícil a sua recuperação (Damme 2006).

Numa linguagem geral, o termo "excesso de informação", ou "*information overload*", é frequentemente utilizado para transmitir a simples noção de receber muitas informações (Eppler and Mengis 2003). Para Eppler e Mengis (Eppler and Mengis 2003), no que diz respeito ao contexto empresarial, as principais causas para o excesso de informação estão relacionadas com a combinação de cinco factores: a própria informação (sua quantidade, frequência ou intensidade, e qualidade ou características gerais), a pessoa que recebe, processa ou comunica a informação, as tarefas ou processos que precisam ser completados por uma pessoa, equipa ou organização, o design organizacional (ou seja, estruturas de trabalho formais e informais), e a tecnologia de informação que é usada (ou como ela é usada).

Existe também uma diferença entre dados estruturados, não-estruturados e semi-estruturados. Dados estruturados são organizados de uma maneira bastante regular, como em tabelas e relacionamentos, onde as regularidades se aplicam a todos os dados num determinado conjunto. Os não-estruturados, como textos ou imagens, contêm informações mas não contêm nenhuma estrutura explícita de informações. No entanto, técnicas manuais ou automáticas podem atribuir *tags* para estas informações, convertendo-as de não-estruturadas para semi-estruturadas (Losee 2006).

Segundo (Damme 2006), existem três importantes factores que justificam a implantação de uma *folksonomy* numa empresa. São eles:

Recuperação de informações não-estruturadas:

Uma vantagem competitiva para as empresas reside consequentemente na habilidade de correlacionar informações estruturadas com informações não-estruturadas, ou semi-estruturadas. No entanto, em alguns casos, isto pode dar origem a problemas relacionados com o acesso da informação (quando são utilizadas tecnologias diferentes para suportar e manipular os diferentes tipos de informações), integridade e disciplina entre os tipos de informações (uma palavra com o mesmo significado nas informações estruturadas e não-estruturadas), e sincronização entre elas (Damme 2006).

Actualmente, a catalogação de informações não-estruturadas é um problema crónico, que se não for adequadamente tratado pode influenciar negativamente o sucesso de uma organização. Além do excesso de informação, a falta de tempo para correlacionar, categorizar, analisar e agir sobre ela é um problema crucial. Apesar das informações estarem na organização, contidas em relatórios, correios electrónicos ou outros documentos, por vezes, quando necessárias para tomadas de decisões rápidas e importantes, elas não são encontradas. Ao contrário das informações estruturadas, elas não podem ser guardadas em linhas e colunas de bases de dados, tais como dados de clientes, transacções, etc.. À medida que o volume de oportunidades aumenta, a duração do tempo de acção para cada oportunidade diminui. Desta forma, uma economia baseada na informação corre grande risco de se afundar num mar de dados irrelevantes e não-estruturados (Delphi 2002).

Para o Delphi Group, os trabalhadores do conhecimento - pessoas que criam informações valiosas para resolver e definir problemas a partir do processamento de informações existentes (Drucker *apud* (Damme 2006)) – em mais de 25% do seu dia estão envolvidos na pesquisa de informações nos seus computadores. Cerca de 70% deste tempo é gasto para navegar nelas (Delphi 2002).

Classificar ou catalogar informações não-estruturadas não é uma tarefa fácil. Além disto, é importante que a informação esteja acessível entre os colaboradores e através dos departamentos, pois isto reduz gastos operacionais (Dirks *apud* (Damme 2006)).

Alto custo da taxonomia e softwares de taxonomia:

Uma taxonomia pode ser criada e mantida por classificadores humanos, por *software*, ou pela combinação de ambos (Delphi 2002). A maior desvantagem da classificação manual é que ela é muito cara e não tem a possibilidade de se adaptar rapidamente às mudanças das

informações. Consistência e eficiência são vantagens de uma classificação automática em relação à classificação manual. Por outro lado, pela falta de controlo e inteligência humana, podem não oferecer alta precisão na classificação das informações. Além disto, *softwares* para uma classificação automática são caros, e muitas vezes inacessíveis às empresas de pequena dimensão. Taxonomias manuais e automáticas são bons mecanismos de classificação, mas ambas possuem a desvantagem de serem muito caras (Damme 2006).

Paralelismo entre informações não-estruturadas nas empresas e informações na Web:

Numa folksonomia para a *Web*, tudo o que possui um endereço URL é elegível. Isto varia de *papers* académicos, artigos, fotos etc., quase os mesmos tipos de informações não-estruturadas (apresentações, documentos etc.) disponíveis e armazenadas nas empresas. Com isto, existe um paralelismo no conteúdo entre as folksonomias para a *Web* e informações não-estruturadas nas empresas (Damme 2006).

## ***4.2 Implementação de uma Folksonomia Empresarial***

Na Web, as folksonomias são um enorme sucesso pelo facto dos utilizadores individuais retirarem benefícios pessoais delas, ou seja, eles recuperam as suas próprias coisas mais facilmente. No entanto, nas empresas algumas questões devem ser consideradas (Damme 2006):

- Como motivar as pessoas para fazerem isto nas empresas?
- Quantos acreditam que este processo poderia ajudá-los a recuperar mais facilmente as suas próprias coisas, bem como reduzir o tempo de pesquisa de informação, optimizando o seu tempo para se poderem dedicar a outras tarefas?
- Isto não beneficiará somente os proprietários ou investidores da empresa?
- Como resolver os problemas de direitos de acesso, devido ao aspecto social da folksonomia na Web, onde qualquer pessoa pode aceder ao conteúdo e marcá-lo com as suas próprias tags?
- Perderá a administração o controlo com este tipo de taxonomia bottom-up, uma vez que a sua qualidade dependerá das entradas de informações dos colaboradores?



A autora acredita que a resposta para estas questões está na cultura da empresa. Baseados em duas dimensões, sociabilidade e solidariedade, (Goffee and Jones 1996) distinguem 4 tipos de culturas de empresas.

Sociabilidade é a medida emocional. São relações onde os indivíduos se consideram mutuamente amigos. Eles tendem a compartilhar certas ideias, atitudes, interesses e valores, normalmente associados em igualdade de condições. Promove o trabalho em equipa, a partilha de informações, um espírito aberto a novas ideias e livre expressão, permitindo assim, um salto para a criatividade. Cria um ambiente no qual os indivíduos estão propensos a ir além dos requisitos formais do seu trabalho (Goffee and Jones 1996).

Por outro lado, a solidariedade é a medida não emocional. São relações baseadas em tarefas ou objectivos comuns, interesses mútuos que irão beneficiar todos os intervenientes. Estas relações podem ser criadas quando surgem as necessidades e desaparecer quando elas forem satisfeitas. A solidariedade gera um elevado grau de foco estratégico, resposta rápida às ameaças competitivas, intolerância e mau desempenho. Constrói um sentido de comunidade na experiência compartilhada, como também desenvolve um forte sentimento de confiança na organização (Goffee and Jones 1996).

Baseados nestes conceitos, Goffee e Jones (Goffee and Jones 1996) apresentam as seguintes culturas corporativas e ressaltam que um *mix* desejável para que uma cultura funcione bem varia de acordo com o contexto de cada empresa:

- **Networked organization [trabalho organizacional em rede] (alta sociabilidade e baixa solidariedade):** É um ambiente mais informal. Os colaboradores são amigos ou foram indicados por amigos. Criaram laços de união dentro dos departamentos e compartilham informações entre si, mas não com outras unidades. Algumas vezes agem como uma família, participando de eventos pessoais dos outros colaboradores. A comunicação entre as divisões é bastante pobre. Esta cultura funciona bem quando estratégias empresariais estão definidas a longo prazo, quando o conhecimento das especificidades dos mercados locais é um factor de sucesso, e quando o sucesso corporativo é um agregado do sucesso local.
- **Mercenary organization [organização mercenária] (baixa sociabilidade e alta solidariedade):** Quase toda a comunicação está centrada nas questões comerciais. Os colaboradores separam bem a vida pessoal da vida profissional, havendo confraternização apenas em torno dos negócios. É caracterizada pela sua capacidade de resposta rápida e coesa a uma ameaça ou oportuni-

dade percebida no mercado. Devido ao elevado foco nos objectivos, os colaboradores não têm tempo para cooperar, distribuir informações, trocar ou criar novas ideias. Cooperação entre unidades com diferentes objectivos é muito menos provável. É funcional quando as mudanças são rápidas e bruscas, quando as vantagens competitivas são obtidas através da criação de centros de excelência empresariais que podem impor processos e procedimentos operacionais em empresas ou divisões, quando os objectivos empresariais são claros e mensuráveis, e quando a natureza da concorrência é clara.

- **Fragmented organization [organização fragmentada] (baixa sociabilidade e baixa solidariedade):** Não há grande envolvimento dos colaboradores com a organização. Eles acreditam que trabalham para si próprios ou identificam-se com grupos ocupacionais, normalmente profissionais. As pessoas trabalham isoladas, com as portas fechadas, ou até mesmo em suas próprias casas. Frequentemente os seus projectos são secretos e somente compartilham informações quando há algum interesse ou retorno. Os membros raramente concordam sobre objectivos organizacionais, factores críticos de sucesso, desempenho e normas. É apropriada quando há pouca dependência do trabalho propriamente dito, quando inovações significativas são produzidas primariamente por indivíduos e não por equipas, quando as normas são atingidas por controlo de entradas e não por controlo de processos, quando há poucas oportunidades de aprendizagem entre os indivíduos ou quando o orgulho profissional impede a transferência de conhecimento. Este tipo de cultura corporativa prevalece em escritórios de consultoria ou escritórios de advocacia, onde cada colaborador tem o seu próprio método de trabalho.
- **Communal organization [organização comunitária] (alta sociabilidade e alta solidariedade):** Os colaboradores possuem um elevado nível de consciência de filiação e identidade organizacional. Relacionam-se eles próprios com a identidade da empresa. Os eventos sociais são organizados para fortalecer a vida organizacional. Riscos e recompensas são igualmente distribuídos entre os colaboradores. Respeito e justiça são muito importantes. Para alguns contextos de negócios, esta cultura pode ser um ideal inadequado e inatingível. É difícil, por exemplo, encontrar empresas comerciais neste quadrante. *Communal organization* pode ser bem aplicada em organizações religiosas, políticas ou civis. Esta cultura resulta bem quando a inovação precisa de uma equipa de trabalho elaborada e extensiva, quando existem reais sinergias entre subunidades organizacionais e verdadeiras oportunidades de aprendizagem, quando as estratégias

são mais a longo-prazo do que a curto-prazo, quando o ambiente empresarial é dinâmico e complexo.

Segundo (Damme 2006), a cultura corporativa mais indicada para a implantação de uma folksonomia é a *Communal Organization*. Para a autora, os colaboradores vão sentir-se-ão altamente identificados com a empresa e definitivamente farão alguns esforços para construir uma taxonomia *bottom-up*. Esta é uma cultura que tem mais oportunidades de ser aceite pelos gestores, pois eles estarão preparados para abrir mão do controlo, porque sabem que os colaboradores farão o seu melhor para colaborar com a inclusão de *tags* para o conteúdo. O grande sentimento de solidariedade entre colaboradores e empresas permitirá que os gestores percebam a importância de autorizar o acesso de todos os colaboradores ao conteúdo, a fim de promover a qualidade das *tags* e, conseqüentemente, o processo de pesquisa.

#### **4.1.1 Vantagens de implementar uma folksonomia empresarial**

As folksonomias poderão promover mudanças fundamentais e importantes nas organizações. Taxonomias rígidas, pouca capacidade de pesquisa, excesso de e-mails etc., são frustrações dos utilizadores que poderão ser minimizadas com o uso destas ferramentas (Cormier 2006).

Um processo de *tagging* pode ser relativamente económico, rápido e fácil de implementar (Hayman 2007).

A actividade de *tagging* colaborativa numa empresa pode permitir a descoberta de novos peritos e a formação de redes sociais de trabalho em torno de *tags* e tópicos (Braly and Froh 2006). A actividade de *tagging* de um colaborador contribui para a sua experiência e influencia na experiência de outros colaboradores (Hayman 2007).

Pelo facto das *tags* serem criadas pelos colaboradores, elas podem ser úteis, facilitando a democracia no trabalho. Além disto, podem também encorajar a colaboração e distribuição de recursos dentro da empresa (Hayman 2007).

A análise dos dados produzidos por estes sistemas podem revelar padrões de utilização de *tags*, o que tornaria possível a sugestão de novos termos e relacionamentos semânticos que incrementariam sistemas e vocabulários controlados já existentes (Braly and Froh 2006).

Oferecem potencialmente uma forma poderosa das empresas organizarem informações recentes, tornando-as imediatamente disponíveis para pesquisa. Além disto, podem revelar-se muito importantes para conhecer a opinião dos colaboradores de diferentes departamentos (ex.: investigação e vendas) e com diferentes ideias sobre detalhes de determinado produ-

to, sendo que alguns estão mais concentrados na lista de materiais e detalhes de fabricação (investigação) e outros em incentivos de vendas e margens de lucro (vendas). A análise de *tags* é outra forma para uma empresa testar conceitos de produtos. Basta permitir que os utilizadores incluam *tags* para o produto desejado e depois analisar os termos que foram usados (Fitzgerald 2007).

*Tagging* pode ser uma maneira de identificar experiências e interesses e fomentar uma maior satisfação e empenho no trabalho dos colaboradores. É também uma forma de captar os seus conhecimentos e avaliações sobre os diversos recursos de informação da *internet* e *intranet* (Hayman 2007).

Poderiam eliminar barreiras geográficas e funcionais entre equipas de trabalho, criando benefícios como a partilha de informações sobre clientes, perspectivas e tendências de mercado, notícias, ideias e recursos de produtos, a reunião de informações sobre consumidores e inteligência competitiva. Possibilitaria aos gestores uma rápida e fácil visão sobre a inteligência colectiva das suas organizações, bem como a conexão existente entre os seus colaboradores (Cook 2006).

Niall Cook (Cook 2007) apresenta quinze utilizações destes sistemas:

- Acompanhar notícias/*blogs* de interesse da empresa;
- Pesquisa de consumidores e concorrentes;
- Identificar peritos em determinados assuntos dentro da empresa;
- Conectar pessoas que tenham os mesmos interesses;
- Verificar o que os seus colegas estão a fazer de interessante agora;
- Inscrever-se em *links* que membros da sua equipa estão a utilizar em *bookmarks*;
- Encontrar as páginas mais populares na *intranet*;
- Utilizar *tags* para classificar conteúdos e complementar pesquisas na *intranet*;
- Verificar o que um pedaço do conteúdo da *intranet* realmente significa para os colaboradores;
- Pesquisar o cérebro colectivo corporativo (em vez de documentos que são "publicados");

- O uso da folksonomia para refinar a taxonomia corporativa;
- Recuperar e partilhar informações sobre clientes e perspectivas;
- Acompanhar tendências da indústria;
- Analisar as conexões entre colaboradores através das equipas e localidades;
- Identificar possíveis contratações e avaliar estudantes.

A folksonomia é uma ferramenta de pesquisa pobre em relação a pesquisas mais eficientes e elaboradas nas *intranets* corporativas, e realmente não interessa para grande parte da empresa, mas o facto é que, em alguns casos, são os utilizadores que fazem a maior parte do trabalho, e tanto eles quanto as empresas obtêm o benefício. Além das *tags* serem sinónimos de palavras-chave familiares para qualquer pessoa que faz uma pesquisa tanto na *internet* quanto na *intranet*, elas servem para qualquer tipo de conteúdo, incluindo vídeo e áudio. A taxonomia permite que todos os colaboradores tenham termos em comum para todos os objectos, mas com *tagging* os colaboradores ganham flexibilidade para trabalhar fora da taxonomia (Fitzgerald 2007).

#### **4.1.2 Desvantagens de implementar uma folksonomia empresarial**

Falta de precisão na escolha dos termos (sem controlo de sinónimos/antónimos, termos relacionados, contextos etc.) (Hayman 2007).

Alguns recursos estão disponíveis somente através de uma conexão de *Internet*. Além disto, *firewalls* e políticas corporativas também podem restringir o seu uso (Cormier 2006).

Informações restritas, mesmo controladas por *firewalls* (os quais impedem o acesso de pessoas de fora da organização), ao serem partilhadas, embora limitadas por natureza, poderiam ser divulgadas (Millen and Feinberg 2006).

Não há garantias de que as pessoas irão incluir *tags* de maneira consistente e imparcial, como também existe a possibilidade de malícia intencional direccionada à empresa (Hayman 2007).

Poderia haver grandes diferenças entre as *tags* já utilizadas na empresa e as *tags* dos novos colaboradores. Uma forma de resolver este problema seria fornecer uma lista com um vocabulário controlado baseado na folksonomia utilizada até ao momento (Damme 2006).

A hierarquia natural da maioria das organizações pode tornar uma classificação *bottom-up* difícil. O carácter informal e igualitário desta abordagem pode não se encaixar bem com uma estrutura decisória *top-down* (Hayman 2007).

Nas empresas com filiais em diversos países as pessoas provavelmente utilizariam diferentes palavras para expressar a mesma coisa. Uma das vantagens da folksonomia é que as comunidades podem ser criadas baseadas no uso de *tags*. A comunicação entre essas diferentes comunidades poderia ser resolvida através da criação de *boundary objects*<sup>2</sup> (Damme 2006).

Determinadas pessoas podem não desejar que outras conheçam os seus interesses e zelar pela sua privacidade. A organização também pode ter preocupações com a visibilidade de uma *tag* num ambiente altamente confidencial (Hayman 2007).

## **4.2 Exemplos de sistemas de organização do conhecimento top-down e bottom-up nas empresas**

Existem diversos exemplos de sistemas de classificação *top-down* e *bottom-up* que estão a ser utilizados nas organizações. Serão apresentados em seguida alguns estudos de empresas que utilizam taxonomias, folksonomias ou uma combinação de ambas.

Em alguns destes estudos, os autores relatam a utilização de *social bookmarking* em vez de folksonomia. *Social bookmarking* é um método para armazenar, organizar e gerir *bookmarks* (marcador de páginas da *Web*) com a ajuda de metadados (Wikipedia 2008), ou seja, com a ajuda de *tags*. Neste caso, devido ao seu carácter social e devido a utilização de *tags* para classificar as URLs, um *social bookmarking* pode ser considerado uma forma de folksonomia.

### **4.2.1 IBM Intranet**

Gibson (Gibson 2005) relata a experiência da IBM na intenção de combinar taxonomia com folksonomia para classificar conteúdos da sua *intranet*, a qual serve aproximadamente 315.000 colaboradores em todo o mundo, em diversos idiomas. A IBM tem usado uma taxonomia controlada com 3.700 nós para organizar esses conteúdos.

---

<sup>2</sup> Boundary Objects (BO), são entidades partilhadas por várias comunidades, mas visualizadas ou utilizadas de maneira diferente por cada uma delas. Eles são flexíveis o suficiente para se adaptarem às necessidades locais e às limitações das diversas comunidades envolvidas (Star, S. L. and J. R. Griesemer (1989). "Institutional ecology, `translations' and boundary objects: amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39." *Social Studies of Science* 19(3): 387-420.).

Essa taxonomia pode ser considerada um vocabulário oficial para abordar assuntos da empresa. No entanto, o grande problema está na quantidade de nós existentes. Aprender sobre 3.700 nós não é uma tarefa fácil para os colaboradores, pois praticamente teriam que aprender um novo vocabulário.

Essa seria uma área onde a folksonomia poderia ajudar. Para (Gibson 2005)

"If people classify items on the intranet with their own free tags in some sort of internal bookmarking system, it would be possible to map between the user's personal vocabulary and the official vocabulary, allowing the user to eventually search with their own vocabulary."

A maioria dos sistemas de *bookmarks* é restritiva por estar vinculada a um navegador específico (Internet Explorer, Firefox etc.). Uma forma de incentivar o uso desse sistema seria desenvolvê-lo desvinculado de qualquer navegador. Um ponto-chave é que o *bookmarking* é fácil e oferece alto valor agregado.

Essas ideias traduziram-se em sistemas, como, por exemplo, os apresentados por (Muller 2007) num estudo comparativo entre quatro serviços da IBM baseados em *tags*:

**TABELA 1 – Comparativo entre serviços da IBM baseados em tags**

Sistemas	Receptor de tags	Número de utilizadores	Número de leitores potenciais
<b>Dogear</b>	URLs públicos, URLs internos e outros documentos internos	1710	Mais de 300.000 colaboradores
<b>Bluepages+1</b>	Pessoas	713	Mais de 300.000 colaboradores
<b>Blog Central</b>	Conteúdos de <i>blogs</i> (somente o autor pode incluir <i>tags</i> )	2092	Mais de 300.000 colaboradores
<b>Activities</b>	Actividades do objecto principal + componentes subordinados das actividades	742	De 0 a várias centenas de colaboradores

Fonte: (Muller 2007)

#### 4.2.2 IBM Dogear (Social Bookmarking)

Devido ao grande sucesso do *social bookmarking* na *Web*, a IBM levantou a questão dos prováveis benefícios que este sistema poderia trazer para as grandes empresas e organizações. Inspirada nesta questão, desenvolveu um sistema interno de *social bookmarking* chamado Dogear. O sistema possui algumas funcionalidades:

- **Identity and transparency (identificação e transparência):** adoção de nomes reais para facilitar a comunicação entre os utilizadores. Isto permitiria inferências sobre interesses e experiências de outros utilizadores, bem como permitiria pesquisar e descoberta de peritos em determinados tópicos para ajudar a resolver problemas de negócios. Facilitaria também a comunicação entre os colaboradores pelo facto deles utilizarem nomes reais nas diversas ferramentas de colaboração corporativa (directórios corporativos, *e-mail*, *chat* etc.). Além disso, possibilita a restrição de *bookmarks* para indivíduos ou grupos, com o objetivo de restringir o acesso a informações confidenciais;
- **Alerting and discovery (alerta e descoberta):** para incentivar o permanente uso de um *social bookmarking*, os utilizadores necessitam ser alertados para novas informações relevantes. Assim, Dogear fornece um *RSS feed* em todas as páginas. O sistema também explora técnicas de *collaborative filtering* para apresentar novos *bookmarks* àqueles que estão interessados nas informações de um indivíduo ou grupo de indivíduos;
- **Designing for extensibility (design abrangente):** incentiva uma integração na medida em que expõe os dados através de uma simples e intuitiva estrutura URL, como também permite aos potenciais exploradores analisá-los facilmente a partir de uma série de formato de dados;
- **Exploiting the links (exploração de ligações):** colecções de *links* podem ser exploradas pela empresa para aprimorar aplicações de pesquisa corporativa, combinar *bookmarks* individuais com outras informações de determinado indivíduo e para partilhar *bookmarks* entre grupos de trabalho.

Como utilizador do Dogear, Bob Zurek (Zurek 2006) da IBM, relata a sua experiência:

"What I find most useful about the Dogear project is how it delivers and surfaces relevant information on demand that my colleagues have tagged while I'm using a search engine. This tagged information, essentially the social bookmark is surfaced right along side my search results. The value of the information from inside and shared by my colleagues has been very valuable."

Num outro artigo, (Millen, Feinberg et al. 2006) os autores fazem uma abordagem mais detalhada sobre o funcionamento do Dogear. Eles apresentam imagens e exemplos do sistema, bem como identificam e relatam cada secção apresentada nas imagens.



### 4.2.3 BluePages+1

Uma das aplicações de maior sucesso na *Intranet* da IBM é um directório empresarial chamado "BluePages", onde os utilizadores pesquisam informações sobre contactos, pessoas da própria organização e contextos sociais. O sistema permite que os utilizadores especifiquem os seus próprios projectos, especializações, experiências, equipas, formação e outros atributos. No entanto, apesar do sucesso do "BluePages", o nível de contribuição individual para os perfis é citado como um problema.

Devido às limitações dos actuais métodos para preenchimento e manutenção dos conteúdos de directórios, a IBM desenvolveu um protótipo chamando "BluePages+1", o qual investiga uma variedade de abordagens para enriquecer e estender as informações disponíveis no directório. Uma delas é integrar automaticamente o conteúdo criado pelos utilizadores noutros sistemas corporativos, tais como *blogs* e *social bookmarks*. O "BluePages+1" tem sido estendido também para utilizadores adicionarem *tags* a outros utilizadores. Cada *tag* adicionada é automaticamente mostrada no conjunto de *tags* daquele perfil. Um dos seus principais objectivos é ajudar os colaboradores a manterem o relacionamento com os seus contactos. Além disto, um potencial benefício do sistema é localizar peritos dentro da empresa.

O sistema foi desenvolvido numa aplicação baseada na *Web* e fundamentado em dois princípios: "*identity and transparency*" e "*low barrier to entry*". O primeiro permite aos utilizadores adicionarem *tags* para outros utilizadores, sem a sua permissão expressa. No entanto, as *tags* são públicas e todos podem ver quem adicionou e quem foi adicionado. O segundo princípio tem por objectivo diminuir os passos requeridos para a adição de *tags*.

No "BluePages+1" cada página de perfil segue um *layout* padrão. O sistema possui as seguintes características:

- **Tag dialog:** espaço disponível em cada perfil destinado à inclusão de *tags*. Tanto os utilizadores que incluem quanto os que recebem uma *tag* tem permissão para a excluir;
- **Tag clouds:** possui duas secções de *tag clouds*. A primeira representa uma colecção de todas as *tags* atribuídas pelos utilizadores ao actual perfil. A segunda secção apresenta a mesma *interface* que a primeira. No entanto, o conteúdo representa as *tags* utilizadas pela própria pessoa do perfil;
- **Contacts icons:** apresenta uma lista de pessoas relacionadas ao perfil actual;

- **Pivoting and searching:** os utilizadores podem pesquisar através das ligações de *tags*. Ao clicar numa *tag*, será apresentada uma lista de utilizadores que foram marcados com ela. O mesmo resultado pode ser obtido ao utilizar palavras-chave num campo de pesquisa encontrado em cada página de perfil;
- **Buddylist import:** permite a importação da lista de amigos do *instant-messenger*.

O sistema prevê a integração com outras aplicações, o que permitirá referenciar pessoas a partir de qualquer lugar que surjam: *e-mail*, *web*, *instant messaging*, documentos etc.

#### 4.2.4 Malibu System

Membros da IBM Watson Research Center (Muller, Geyer et al. 2007) apresentaram um sistema experimental chamado Malibu. Ele foi desenvolvido para auxiliar os trabalhadores nas suas actividades, com o objectivo de melhorar e superar as limitações das actuais abordagens das actividades de gestão.

O sistema funciona como uma barra lateral na área de trabalho (semelhante ao Google Desktop) e possui os seguintes recursos: controlo de tarefas, actividades, *bookmarks* partilhados, *feeds* e indivíduos. Além disso, permite que novos recursos sejam acrescentados. O seu componente chave para suportar pesquisa baseada em *tags* é um navegador localizado no topo da barra lateral. A partir deste navegador é possível efectuar pesquisas em qualquer item de qualquer recurso de dados.

Nas entrevistas realizadas, os utilizadores relataram que Malibu é bastante útil para organizar as suas tarefas e actividades. Por outro lado, relataram também problemas. Para eles, alguns recursos não foram completamente caracterizados. Além disso, geraram expectativas que não foram cumpridas, apesar de terem sido úteis para os utilizadores.

Malibu permite a utilização de informações de *social tag* para combinar diferentes recursos de dados e, ao fazê-lo, pode auxiliar os colaboradores na localização de informações relevantes para melhor apoiar as suas tarefas.

## 4.2.5 Lucent Technologies

Angeles (Angeles 2005) apresenta um caso específico do seu grupo de trabalho na Lucent Technologies<sup>3</sup>. Eles estão a testar uma aplicação de auto-serviço de *publishing*<sup>4</sup>. A ideia de utilizar o conceito de *social bookmarking* e torná-lo mais do que apenas um serviço para guardar *bookmarks* surgiu a partir de vários pedidos dos colaboradores. Um destes pedidos recaía sobre a forma de marcar ou guardar artigos encontrados no *site* da biblioteca digital. Um segundo, e considerado mais importante, foi para a definição de uma forma de os utilizadores guardarem artigos e publicá-los em *portlets*<sup>5</sup> no portal corporativo.

O conjunto de requisitos para o sistema tratou dois temas: *tagging* de registos individuais e reutilização de registos fora do *site*, sendo que o primeiro fornece um meio para marcar os registos e o segundo proporciona uma forma de reutilização dos dados noutros lugares.

O sistema fornece um recurso para salvar, incluir *tags* e comentar uma página *web*, como também integra recursos de XML e RSS *feed* já existentes nas bases de dados da empresa. O sistema possui três características principais:

- **Self-service publishing:** permite que indivíduos ou grupos criem *bookmarks* dos artigos da base de dados de notícias. As *tags* que dão acesso aos artigos podem ser exibidas num *portlet*, facilitando a recuperação das informações;
- **Feeds and exporting:** permite uma maneira mais simples de reutilização de RSS, através de uma saída de HTML juntamente com RSS. Permite também o uso de JavaScript para colocar entradas de *bookmarks* em JS *feed* com as últimas actualizações guardadas numa matriz;
- **Controlling the sprawl:** as funcionalidades do sistema são fraccionadas e tratadas individualmente. Cada fracção é tratada minuciosamente, sendo melhorada e simplificada do ponto de vista do utilizador. Para eles, trabalhar com cenários pequenos e bem definidos é uma tarefa muito mais fácil do que trabalhar com grandes cenários e regras. No final, estes fragmentos integram-se num sistema maior no *site* da empresa, tornando-o mais simples, seguro e consistente.

---

<sup>3</sup> Empresa que fornece soluções que permitam aos prestadores de serviços, empresas e governos em todo o mundo, disponibilizarem serviços de voz, dados e vídeo para utilizadores finais Alcatel-Lucent. (2008). "Alcatel-Lucent." Retrieved March 11th, 2008, from <http://www.alcatel-lucent.com..>

<sup>4</sup> Processo de produção e disseminação de literatura ou informação Wikipedia. (2008). "Publishing." from <http://en.wikipedia.org/wiki/Publishing>.

<sup>5</sup> Portlets são pequenas janelas de conteúdo html que agem como pequenos blocos ou módulos numa página de um portal Angeles, M. (2005). Making libraries more delicious: social bookmarking in the enterprise..

#### 4.2.6 AskMe Corporation

Steve Eisner (Eisner 2006) compartilha a experiência da AskMe Corporation<sup>6</sup> na incorporação de *tagging* no seu produto.

Inicialmente, a aplicação permitia aos editores escolher uma categoria em que o novo conteúdo iria aparecer. Cada item deveria ser publicado somente numa categoria simples, como numa estrutura de directório. Após várias reclamações dos clientes alegando um sistema muito restritivo, a empresa acrescentou a capacidade de incluir conteúdos em múltiplas categorias. No entanto, os seus aspectos hierárquicos foram mantidos, podendo haver campos pais e campos filhos, mas ganhou a possibilidade de sobreposição entre eles.

Apesar do sistema ter os mesmos benefícios de uma *social tagging*, pois permite a participação dos utilizadores na classificação de conteúdos e a navegação através de *tags*, o autor não o considera uma folksonomia, e sim uma "*semi-tagging*". O facto é que apesar da classificação ser livre, as categorias são pré-definidas e não permitem o uso do vocabulário dos utilizadores.

A liberdade de classificação agradou bastante aos editores e uma grande quantidade de conteúdos começou a aparecer em várias categorias. No entanto, além de enriquecer os conteúdos apresentados por cada categoria, bem como facilitar o relacionamento entre elas, esta implementação particular teve alguns aspectos negativos em relação à participação e à segurança:

- **Participação:** naturalmente os leitores preferem incluir *tags* com os seus próprios termos. A obrigatoriedade do uso de termos pré-definidos tornou a participação mais difícil. Assim, comparado aos sistemas que permitem *tags* livres, o aparecimento da "inteligência colectiva" da rede de trabalho tornou-se mais lento;
- **Segurança:** as categorias poderiam ter definições de permissão. No entanto, se um conteúdo fosse restrito mas também classificado em alguma categoria não restrita, ele poderia ser acedido por utilizadores não autorizados. Para resolver este problema foi introduzido um conjunto de regras de permissões.

A partir destes resultados, algumas lições foram apreendidas:

---

<sup>6</sup> A AskMe Corporation é uma empresa de software e serviços profissionais para implementar redes de conhecimento dos colaboradores Realcom, A. (2008). "AskMe Realcom." Retrieved March 12th, 2008, from <http://www.askmecorp.com/company/default.asp>.

- Não deveriam ter usado apenas *tagging*. Ainda não há evidências reais de que uma simples *tagging* conduz a adopção generalizada da empresa;
- Uma forma livre de folksonomia pode não satisfazer as necessidades dos seus clientes;
- As expectativas criadas pela terminologia da *tagging* importa tanto quanto a sua implementação;
- Para a nova versão do produto, decidiram que é melhor separar a terminologia entre actividades de *tagging* pessoais e públicas. Eles acreditam que as pessoas utilizarão *tags* para o seu próprio benefício, mas que poderão contar com rotinas semi-automáticas para tornar os conteúdos públicos.

#### 4.2.7 STARLab: The Guide

Stijn Christiaens (Christiaens 2006) apresenta o caso do portal da comunidade do Semantics Technology and Applications Research Laboratory (STARLab<sup>7</sup>) chamado "The Guide". Os membros da comunidade são pessoas que trabalham no laboratório. O "The Guide" é suportado por Drupal<sup>8</sup>, um sistema *open source* para gestão de conteúdos. O portal é utilizado para gerir todo tipo de informações úteis, que vão desde questões técnicas até profundas discussões teóricas. Alguns conteúdos são totalmente livres enquanto outros devem ser colocados em locais específicos.

Para a classificação de conteúdos é utilizada uma folksonomia restrita, ou seja, as *tags* são atribuídas pelos próprios autores. Elas podem ser visualizadas numa *tag cloud*. Além disto, desenvolveram uma taxonomia derivada da *tag cloud*, através do agrupamento e da ordenação das *tags* mais genéricas para as mais específicas. As mais genéricas tornam-se rótulos dos grupos. O resultado final é uma taxonomia básica que traz mais estrutura ao "The Guide".

Para tirar melhor partido dos conteúdos, desenvolveram uma abordagem para a construção de uma ontologia para o portal. Ela captura os significados das relações entre todos os objectos de informações e descreve todo o conteúdo do sistema.

Após enriquecer o "The Guide" com uma taxonomia (para estruturar as *tags* criadas pelos utilizadores) e uma ontologia (para descrever a própria estrutura do portal), eles procuraram

---

<sup>7</sup> <http://www.starlab.vub.ac.be>

<sup>8</sup> <http://www.drupal.org>

formas úteis de aplicação. Uma mais valia foi possibilitar a construção de grupos de peritos, conforme o exemplo a seguir citado.

Suponhamos que um novo investigador começa a trabalhar no STARLab e gostaria de saber tudo sobre trabalhos anteriores, actuais e futuros. Ele acede à página *Expert Finger* para procurar pessoas que trabalham com determinado assunto. No campo de pesquisa ele digita "desenvolvimento", pois quer dialogar com pessoas que trabalham com ferramentas de desenvolvimento. O sistema procura por todos os *posts* que possuem a *tag* "desenvolvimento". Para incrementar o resultado, o sistema ainda adiciona automaticamente na pesquisa as *tags* que estão categorizadas abaixo de "desenvolvimento" na taxonomia. Depois ele identifica as pessoas que criam ou comentam os *posts* e classifica-as num *ranking* de peritos de acordo com as suas actividades nos *posts*. Ao mesmo tempo o sistema cria um novo tópico no Fórum, dizendo que tem um novo investigador a procurar por informações sobre "desenvolvimento" no STARLab. Em seguida, convida por e-mail as duas pessoas melhores classificadas no *ranking* para discutir questões de "desenvolvimento" com ele. O novo investigador tem autonomia para encerrar o tópico quando assim o entender.

Através da combinação de folksonomia, taxonomia e ontologia eles obtiveram como resultado final um sistema que fornece uma série de possibilidades para gerir e incrementar o conhecimento partilhado entre os seus colaboradores.

## **5 DIFERENÇAS ENTRE OS SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO**

No seguimento dos capítulos anteriores, pretende-se aqui apresentar uma comparação entre os SOC. Mais especificamente estabelecer uma comparação entre a folksonomia – uma maneira informal de organização – com alguns modelos formais, tais como o esquema de classificação universal, facetado, ontologia, taxonomia e tesauro.

### ***5.1 Alguns Estudos Anteriores***

Alguns estudos anteriores foram realizados sobre os SOC, como por exemplo (Beck and Pinto 2002) e (Vickery 2008) com referências aos tesauros e às ontologias; (van Rees 2003) na comparação entre classificação, taxonomia e ontologia; e (Damme 2006) no confronto entre taxonomia e folksonomia.

#### **5.1.1 Tesauro vs Ontologia**

Segundo (Beck and Pinto 2002), o tesauro e a ontologia são semelhantes em vários aspectos. Para os autores, ambos se preocupam com a cobertura de uma vasta gama de terminologias utilizadas num domínio particular, e na compreensão da relação entre estes termos. Ambos utilizam uma organização hierárquica para agrupar os termos em categorias e subcategorias, como também podem ser aplicados numa catalogação e classificação de informações.

No entanto, embora semelhantes numa estrutura geral, a ontologia fornece mais detalhes e representa melhor o conhecimento formal, pois permite introduzir uma série de relacionamentos estruturais e conceptuais, como super-classes, subclasses, exemplos, valores de propriedades, relacionamento de tempo etc., possibilitando um significado mais completo dos termos.

As principais características da ontologia, comparada a outros tradicionais SOC, são que todos os relacionamentos semânticos entre os termos são expressamente determinados, não existem limites para a variedade de relações que pode ser utilizada, e as regras de inferência ligam as afirmações de modo que deduções possam ser feitas a partir de afirmações explícitas para outras que estão logicamente nelas implícitas (Vickery 2008).

No caso do tesauro, não é possível definir conceitos, exemplos ou quaisquer outros valores, a não ser os relacionamentos entre os termos, determinados pelos símbolos BT (*broad term* ou

termo geral), NT (*narrow term* ou termo específico), RT (*related term* ou termo associado ou relacionado) e SY (*synonym* ou sinónimo). Tais símbolos determinam apenas a hierarquia entre cada elemento do tesouro. Para (Vickery 2008), eles não reflectem as relações conceptuais que as pessoas conhecem e que poderiam ser utilizadas por um sistema para sugerir conceitos para ampliar ou tornar mais específica uma pesquisa.

Ainda segundo (Vickery 2008), o tesouro possui algumas limitações. Para ele, a distinção entre um conceito (significado) e o seu léxico (palavra) não é feita maneira consistente, e que nem sempre reflecte a forma como os humanos compreendem o mundo em termos de significado e linguagem. Para além disso, a falta de precisão semântica na estrutura do tesouro pode criar ambiguidade na interpretação dos relacionamentos e resultar numa semântica interna geral irregular e imprevisível.

Em geral, uma ontologia contém muito mais relacionamentos do que um tesouro. Enquanto um tesouro apresenta os relacionamentos entre os termos, uma ontologia apresenta os conceitos e os relacionamentos entre eles. (Beck and Pinto 2002).

Na prática, uma ontologia é melhor aplicável quando se trata de pesquisas. Pelo facto dela possuir mecanismos para interpretação de conceitos, é capaz de sustentar um raciocínio terminológico. Isto significa que as questões dos utilizadores podem ser entendidas através da análise dos significados dos termos presentes nas questões e mapeadas mais precisamente para os recursos de informação. A ontologia pode justificar o significado de conceitos através da comparação lógica da sua estrutura (Beck and Pinto 2002).

### **5.1.2 Classificação vs Taxonomia vs Ontologia**

Num artigo chamado “Clarity in the usage of the terms ontology, taxonomy and classification” (van Rees 2003), podemos encontrar uma comparação entre ontologia, taxonomia e classificação.

Para o autor, a diferença entre uma classificação e uma taxonomia é que a taxonomia classifica uma estrutura de acordo com alguns relacionamentos entre as entidades, enquanto a classificação utiliza razões mais arbitrárias, ou externas às entidades. Por exemplo, o homem é um mamífero, mas nem todo o mamífero é um homem, então o homem é uma subclasse dos mamíferos. Neste caso, a decisão de alocar o homem na categoria dos mamíferos foi baseada em dados inerentes à entidade, o que poderia também ser considerado um elemento de uma taxonomia. Enquanto que a taxonomia possui relações entre as entidades, a classificação relaciona entidades a algo que é externo, como por exemplo ramos de uma indústria.



O autor considera a ontologia um tipo de taxonomia com definições e um tipo de linguagem de representação do conhecimento. Para ele, distinguir entre uma ontologia e uma taxonomia é algo não muito claro, pois da mesma forma que propriedades extras podem ser adicionadas à taxonomia, como uma definição da hierarquia escolhida, uma ontologia pode conter relações taxonómicas.

No caso da classificação e da ontologia, a diferença principal advém da riqueza de informações disponíveis. Ambas proporcionam uma lista ou estrutura de conceitos ou itens de classificação. No entanto, enquanto a classificação pára neste ponto, a ontologia fornece ainda diversas informações sobre conceitos e os seus relacionamentos (van Rees 2003).

Além de comparar estes sistemas de classificação, o autor ainda ressalta que um tesouro, devido às suas características de possibilitar alternativas, sinónimos, interpretações etc. às palavras, poderia ser uma boa contribuição para cada um deles.

### **5.1.3 Taxonomia vs Folksonomia**

Na taxonomia, cada animal, livro, arquivo ou qualquer outro objecto, está numa categoria específica, que por sua vez está dentro de uma categoria mais geral. É o que ocorre, por exemplo, em sistemas de classificação de bibliotecas (ex.: *Dewel Decimal Classification - DDC e Library of Congress Classification - LCC*), onde cada livro tem uma categoria específica dentro de uma categoria mais genérica, como também possui uma posição física exclusiva, mesmo que aborde diferentes assuntos. Enquanto a taxonomia é uma abordagem *top-down*, ou seja, fornece uma determinada hierarquia e um vocabulário de domínio específico, a folksonomia é o oposto, uma abordagem *bottom-up* que não mantém nenhum vocabulário controlado e nenhuma hierarquia explícita (Golder and Huberman 2006); (Al-Khalifa 2007); (Noruzi 2007).

Uma taxonomia restrita força os utilizadores a ter uma visão dos termos de acordo com o ponto de vista do seu criador, que na maioria das vezes não é familiar. Por outro lado, oferece suporte para pesquisa e navegação, ao contrário da folksonomia, que oferece apenas assistência básica. No entanto, a folksonomia permite que os utilizadores escolham as suas próprias palavras-chave para descrever um recurso de uma maneira que faça sentido para eles (Bielenberg and Zacher 2005).

Para (Butterfield 2004), um dos criadores do Flickr, as folksonomias possuem 90% de uma boa taxonomia, mas 10 vezes mais simples. Para ele, a falta de hierarquia, controlo de sinónimos e precisão semântica são precisamente os motivos pelos quais elas funcionam.

O fundador do Delicious, Josua Schachter, tem uma abordagem mais moderada, ressaltando que um esquema de categorização individual vale menos do que um esquema de categorização profissional. No entanto, se as pessoas encontrarem uma maneira útil para incluir *tags* aos seus objectos, tal facto irá gerar muito mais dados do que se o serviço fosse contratado a um profissional especializado para os classificar apenas uma vez. Caso seja encontrada uma maneira de agregar valor para este sistema amador de classificação, ele tornar-se-á mais valioso do que os sistemas profissionais de classificação, em particular no que se refere à robustez e custos de criação (Shirky 2005).

Segundo (Quintarelli 2005), as folksonomias não são a solução para todos os problemas actuais de classificação, como também não são alternativas para os tradicionais esquemas de classificação. Para a autora, a folksonomia é uma poderosa e inovadora ferramenta que deve ser aplicada apenas nas circunstâncias certas, considerando as suas próprias características e as diferenças em relação aos sistemas formais de classificação.

Quintarelli distingue potenciais áreas de aplicação para taxonomia ou folksonomia:

- **Taxonomias:** são adequadas para a classificação de entidades homogéneas, estáveis, entidades restritas com uma entidade central e utilizadores qualificados. Por outro lado têm alto custo de criação e manutenção;
- **Folksonomias:** exigem pessoas para fazer o trabalho, tanto por razões pessoais quanto por razões sociais. Não possuem hierarquia, são ambíguas e não suportam uma pesquisa assistida. No entanto, ao contrário das taxonomias, as folksonomias têm baixo custo, são escaláveis e aproximam-se da linguagem e modelo mental dos utilizadores.

Algumas características destas abordagens são apresentadas por (Damme 2006):

#### QUADRO 1 – Folksonomia vs Taxonomia

<b>Folksonomia</b>	<b>Taxonomia</b>
Livre escolha de palavras-chave	Vocabulário controlado
Sem uma autoridade central	Autoridade central
Baixo custo	Alto custo
Imprevisível	Previsível

Fonte: (Damme 2006)

As folksonomias são conjuntos emergentes de *tags* sem uma organização hierárquica estruturada, criadas por utilizadores ao mesmo tempo que se publica, insere ou cataloga itens. Em contraste, as taxonomias são normalmente concebidas antes de iniciar o catálogo de itens. Elas são elaboradas por profissionais através de previsões das necessidades dos utilizadores e tipologias de conteúdos. As taxonomias propõem uma visão autoritária central, enquanto as folksonomias deixam emergir uma visão colaborativa descentralizada. As taxonomias possuem alta precisão, evitam ambiguidade e a sua estrutura hierárquica contribui para dar contexto aos termos. Por outro lado, sistemas de *tagging* são imprecisos por definição e actualmente não oferecem controlo de sinónimos (Quintarelli 2005); (Guy and Tonkin 2006).

As folksonomias são claramente atraentes, suportam uma forma de pesquisa imprevisível, com descobertas inesperadas, que em muitos casos se revelam úteis. Mas elas não suportam pesquisas avançadas e outras formas de navegação, bem como *tags* de vocabulários controlados aplicados por profissionais (Rosenfeld 2005). Para Rosenfeld, as folksonomias podem trabalhar bem para arquivos de fotos, por exemplo, mas em áreas mais especializadas, como na medicina, um vocabulário controlado ainda é fundamental. Para ele, neste caso, a folksonomia poderia ser uma ferramenta auxiliar para incrementar os resultados da pesquisa (comentário em (Shirky 2005)).

Shirky (Shirky 2005) afirma que

"The advantage of folksonomies isn't that they're better than controlled vocabularies, it's that they're better than nothing, because controlled vocabularies are not extensible to the majority of cases where tagging is needed. Building, maintaining, and enforcing a controlled vocabulary is, relative to folksonomies, enormously expensive, both in the development time, and in the cost to the user, especially the amateur user, in using the system."

Céline Van Damme considera ousado dizer que a folksonomia pode ser uma substituta para a taxonomia, apesar de ela ser muito valiosa. Através de uma análise SWOT<sup>9</sup>, (Damme 2006) identificou vantagens, desvantagens, oportunidades e ameaças de uma taxonomia e uma folksonomia:

---

<sup>9</sup> O termo SWOT resulta da conjugação das iniciais das palavras anglo-saxónicas Strengths (forças), Weaknesses (fraquezas), Opportunities (oportunidades) e Threats (ameaças). Assim, a análise SWOT corresponde à identificação por parte de uma organização e de forma integrada dos principais aspectos que caracterizam a sua posição estratégica num determinado momento, tanto a nível interno como externo (forma como a organização se relaciona com o seu meio envolvente). Infopédia. (2003-2008, September 14th, 2008). "análise SWOT." Infopédia, from [http://www.infopedia.pt/\\$analise-swt](http://www.infopedia.pt/$analise-swt)..

**QUADRO 2 – Análise SWOT: Folksonomia e Taxonomia**

	<b>Folksonomia</b>	<b>Taxonomia</b>
<b>Vantagens</b>	<p>Não hierárquica;</p> <p>Fácil de manter;</p> <p>Sem autoridade central que impõe um determinado vocabulário. O que predomina é o vocabulário dos utilizadores;</p> <p>Imprevisível. Permite encontrar coisas numa área inesperada;</p> <p>O mecanismo de <i>feedback</i> pode ser visto como um tipo de mecanismo de sinónimos;</p> <p>Vantagem perante a não aplicabilidade de uma taxonomia;</p> <p><i>Tags</i> como elos de ligação entre indivíduos.</p>	<p>Devido à estrutura hierárquica, é muito fácil navegar para obter informações;</p> <p>Possui uma autoridade central para controlo, o que diminui a proliferação de categorias;</p> <p>Vocabulário controlado, com variantes e termos preferidos;</p> <p>Categorias bem definidas.</p>
<b>Desvantagens</b>	<p>Problemas com polissemias, sinónimos e variações de níveis básicos;</p> <p>Aumento de <i>tags</i> personalizadas, o que pode dificultar a recuperação de conteúdos;</p> <p>Falta de precisão por não haver controlo de sinónimos;</p> <p>Na maioria das vezes o espaçamento não é permitido nas <i>tags</i>;</p> <p>Não há controlo de vocabulário;</p> <p><i>Tags</i> não possuem uma boa organização pela falta de hierarquia.</p>	<p>Problemas com polissemias, sinónimos e variações de níveis básicos;</p> <p>Baixa compreensão e rigidez do vocabulário controlado;</p> <p>"Poluição" devido ao vocabulário controlado. Ao classificar documentos numa categoria errada, os utilizadores acabam por "poluir" o sistema;</p> <p>Ao classificar um documento, a hierarquia obriga a que seja escolhida uma categoria mais importante do que as outras;</p> <p>Trabalhosa e com alto custo associado.</p>
<b>Oportunidades</b>	<p>Implementação de regras para <i>tagging</i>;</p> <p>Verificador de grafias;</p> <p><i>Ranking</i> de utilizadores;</p> <p><i>Clustering</i>, ou seja, uma maneira de determinar quais as <i>tags</i> mais importantes num conjunto de <i>tags</i>;</p> <p>Implementação de uma hierarquia.</p>	<p>Integração de uma folksonomia com uma taxonomia.</p>
<b>Ameaças</b>	<p><i>Tags</i> personalizadas ou "egoístas" podem tornar-se uma ameaça real para o valor da folksonomia.</p>	<p>Excesso de nós;</p> <p>Polihierarquias, ou seja, termos que podem ser classificados em mais de um super-tipo;</p> <p>Hierarquia muito restrita e densa ou muito abrangente e superficial.</p>

Fonte: (Damme 2006)

Em alguns casos, um sistema baseado em *tagging* pode ter vantagens sobre uma taxonomia. Por exemplo, caso seja feito um *download* de um artigo sobre "espécies nativas de gatos da África" e queira organizá-lo segundo uma taxonomia, poderiam ser observadas algumas opções, tais como (Golder and Huberman 2006):

- c:\artigos\gatos (todos os artigos sobre gatos);
- c:\artigos\afrika (todos os artigos sobre África);
- c:\artigos\afrika\gatos (todos os artigos sobre gatos africanos);
- c:\artigos\gatos\afrika (todos os artigos sobre gatos de África).

Para os autores, neste caso, cada escolha reflecte uma decisão sobre a importância de cada categoria. Algumas limitações podem ser observadas, como no caso das duas primeiras opções, onde os ficheiros representam de facto "gatos" ou "África", mas não permitem uma classificação mais específica para o artigo. Nas duas últimas opções, os ficheiros são organizados por ambas as categorias, no entanto, estabelecem a primeira como a mais geral. A decisão a ser tomada pelo utilizador implicará todas as futuras classificações, inclusive na possível duplicação de arquivos. Apesar destas limitações, há boas razões para implementar um sistema baseado numa taxonomia. Manter alguns ficheiros numa determinada hierarquia pode ser uma forma eficiente e específica, ou não ambígua, de classificar conteúdos. Assim, o utilizador pode estar seguro de que todos os arquivos sobre determinado assunto estarão num mesmo local.

Em contraste, um sistema de *tagging* permitiria identificar o mesmo artigo com uma grande variedade de palavras-chave, sem qualquer restrição, sendo apenas necessário marcá-lo como conveniente. Neste caso, o utilizador poderia não ter a certeza de que nos resultados de uma pesquisa estariam todos os arquivos relacionados com aquele assunto. Por exemplo, se para um arquivo fosse dada a palavra-chave "leopardo", ele certamente não apareceria nos resultados de uma pesquisa feita com as palavras "gato" e "África". Desta forma, talvez o utilizador necessitasse de aprofundar a sua pesquisa para chegar a um resultado mais preciso. Por outro lado, em nenhum momento o documento precisaria ser duplicado para constar em diferentes categorias, como pode ocorrer numa taxonomia (Golder and Huberman 2006). Em sistemas baseados em *tagging*, não existe nenhuma restrição em relação à localização física dos objectos. É perfeitamente possível que *links* estejam disponíveis em diferentes lugares numa hierarquia ou em diversas hierarquias ao mesmo tempo (Shirky 2005).

Apesar das diversas discussões sobre folksonomia e taxonomia, onde em determinados momentos elas são colocadas em lados opostos, é importante ressaltar que estas duas abor-

dagens estão relacionadas à organização e recuperação de informações. Se por um lado a taxonomia oferece um melhor controlo e organização das informações com atributos bem definidos, como títulos, autores e assuntos de livros ou artigos, a folksonomia ajusta-se melhor quando se trata de objectos com valores mais específicos e não tão claros, como, por exemplo, objectos não textuais (Fitzgerald 2007).

Ao classificar um livro ilustrado, por exemplo, na indexação do título, autores, assunto e índice, a taxonomia é a que melhor se adequa, pois estes são atributos claros, estruturados e comuns a qualquer indivíduo. No entanto, caso seja necessário classificar as ilustrações contidas neste livro, a indexação torna-se mais difícil, pois, neste caso, ela está associada à interpretação do classificador. Desta forma, devido à sua flexibilidade e capacidade em incorporar novos termos, a folksonomia é a melhor opção. Além disso, ela permite que indivíduos com interesses nos mesmos conceitos e/ou termos partilhem da mesma classificação e dos mesmos conteúdos.

Em síntese, ao mesmo tempo que estes dois sistemas de classificação possuem características muitas vezes opostas, eles podem ser utilizados como complemento um do outro, com o objectivo de indexar o maior número de informações relacionadas com um objecto.

## ***5.2 Folksonomia vs Sistemas Formais de Classificação***

Nesta secção estabelece-se uma comparação entre um sistema informal e sistemas formais de classificação. A análise aqui apresentada não possui nenhum cunho prático. O que se segue é um confronto entre as características encontradas na literatura de cada sistema em questão.

Não existe nenhuma metodologia consensual para avaliar sistemas de classificação. No entanto, com base no estudo de (Pomerantz 2003), onde são apresentados treze critérios para avaliar esquemas de classificação, e nos aspectos relevantes de cada um dos até agora caracterizados, elaborou-se uma grelha de comparação com vinte e um critérios divididos em cinco categorias, conforme descrito no quadro seguinte. Este conjunto de critérios não é suficientemente específico para dar uma visão detalhada dos aspectos mais importantes de cada sistema, mas possibilita a sua comparação a partir de uma perspectiva mais geral, ou seja, de uma perspectiva do sistema como um todo.

**QUADRO 3 – Critérios para Avaliação dos Sistemas de Classificação**

Categorias / Critérios		Descrição
<b>Estrutura</b>	Escopo	Permite a cobertura de todas as áreas temáticas.
	Holística	O esquema de classificação como um todo fornece uma visualização do fenómeno que ele está a representar.
	Expressividade	Permite a incorporação da estrutura ou vocabulário que melhor atende a representação do conhecimento, ou seja, a naturalidade da linguagem na classificação das entidades.
	Visualização	Ao visualizar a estrutura, as entidades e os seus relacionamentos estão evidentes.
<b>Classes</b>	Hierarquia	Define a hierarquia entre as entidades.
	Granularidade	Consistente nível de especificidade das classes, ou seja, as categorias estão bem definidas.
	Hospitalidade	Fácil acomodação de novas classes ou entidades que não se encaixam na estrutura existente.
	Actualização	Permite rápida e fácil actualização face às mudanças nas tecnologias e aos eventos mundiais.
	Hereditariedade	As informações podem ser transportadas através da estrutura.
<b>Terminologia</b>	Coerência	Lógica e unidade no vocabulário e na estrutura. Pouca ou nenhuma subjectividade na classificação que possa levar a uma falta de coerência.
	Consistência	Previsibilidade do vocabulário e estrutura.
	Multilingue	Permite uma classificação multilingue.
	Terminologia controlada	Os termos são pré-estabelecidos por especialistas ou grupos de interesse, não permitindo qualquer tipo de linguagem externa àquela.
	Vocabulário	Profundidade e consistência dos termos para referenciar as classes.
<b>Utilidade</b>	Usabilidade	Fácil utilização na perspectiva do classificador.
	Navegabilidade	Fácil utilização na perspectiva do utilizador.
	Serendipidade	Permite descobertas ao acaso.
	Democrático / Colaborativo	Permite que qualquer pessoa altere ou sugira alterações.
	Comunitário	Permite identificar grupos de interesses comuns através da classificação.
<b>Custos</b>	Custo económico	Custo financeiro para implementação e manutenção.
	Dispêndio de tempo	Dispêndio de tempo na implementação e manutenção.

Mesmo em contextos semelhantes, as organizações são bem diferentes e possuem necessidades distintas. Escolher um sistema para a gestão do conhecimento é uma tarefa importante e

delicada, pois uma má escolha pode influenciar nos objectivos ou resultados da organização. Por outro lado, as empresas nem sempre podem optar pelo sistema que mais se adapta às suas necessidades, tendo muitas vezes que conjugar este factor com outros. Os custos envolvidos, por exemplo, podem ser determinantes para a escolha entre os vários sistemas existentes. Desta forma, agrupar os critérios de avaliação acima descritos em categorias permite aferir em cada conjunto o seu valor e relevância.

De acordo com as suas particularidades, os critérios foram avaliados a partir de uma escala de valores contidos entre 0 e 1, distribuídos da seguinte maneira: nulo (0); baixo (0,33); médio (0,67); alto (1). O que se pretendeu com esta avaliação foi identificar em que critérios os sistemas de classificação eram comuns ou divergentes. A partir desta perspectiva, conforme demonstrado no quadro infra, foi possível confrontar a folksonomia com o esquema de classificação universal, facetado, tesauro, ontologia e taxonomia.

**QUADRO 4 – Comparação entre os Sistemas de Classificação**

Critérios de Comparação		Sistemas de Classificação					
		Folksonomia	Universal	Facetado	Tesauro	Ontologia	Taxonomia
Estrutura	Escopo	1	1	1	1	1	1
	Holística	0,33	0,33	0,33	0,33	1	0,67
	Expressividade	1	0,33	1	0,67	0,67	0,33
	Visualização	0	1	0	0,33	1	1
	<b>Sub-Total</b>	<b>2,33</b>	<b>2,66</b>	<b>2,33</b>	<b>2,33</b>	<b>3,67</b>	<b>3</b>
Classes	Hierarquia	0	0,33	0	1	1	1
	Granularidade	0,33	1	0,33	0,67	1	1
	Hospitalidade	1	0,33	1	0,33	0,67	0,67
	Actualização	1	0,33	1	0,33	0,67	0,67
	Hereditariedade	0	0,67	0	0,33	1	1
	<b>Sub-Total</b>	<b>2,33</b>	<b>2,66</b>	<b>2,33</b>	<b>2,66</b>	<b>4,34</b>	<b>4,34</b>

Continua...



...Continuação

Terminologia	Coerência	0,33	1	0,33	0,67	1	1
	Consistência	0,33	1	0,33	0,67	1	1
	Multilingue	1	1	1	1	1	1
	Terminologia controlada	0	1	0,33	1	1	1
	Vocabulário	0,33	1	0,67	1	1	1
	<b>Sub-Total</b>	<b>1,99</b>	<b>5</b>	<b>2,66</b>	<b>4,34</b>	<b>5</b>	<b>5</b>
Utilidade	Usabilidade	1	0,33	1	0,33	0,33	0,33
	Navegabilidade	1	0,33	1	0,67	0,67	0,67
	Serendipidade	1	0	0	0	0	0
	Democrático/Colaborativo	1	0	0	0	0	0
	Comunitário	1	0	0	0	0	0
	<b>Sub-Total</b>	<b>5</b>	<b>0,66</b>	<b>2</b>	<b>1</b>	<b>1</b>	<b>1</b>
Custos	Baixo Custo económico	1	0	0	0	0	0
	Baixo Custo de tempo	1	0	0	0	0	0
	<b>Sub-Total</b>	<b>2</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>
<b>TOTAL</b>		<b>13,65</b>	<b>10,98</b>	<b>9,32</b>	<b>10,33</b>	<b>14,01</b>	<b>13,34</b>

Inicialmente, ao analisar os sistemas de classificação do ponto de vista das categorias de critérios, pode-se observar que em relação à categoria “Estrutura” a ontologia é a que mais se destaca, seguida pela taxonomia e esquema universal. Por outro lado, com uma avaliação menos elevada estão a folksonomia, sistema facetado e tesouro.

Nas categorias “Classe” e “Terminologia”, a tendência é a mesma que na categoria anterior, tendo a ontologia, taxonomia e esquema universal com os valores mais elevados, e a folksonomia e o sistema facetado com os menores valores.

No entanto, nas categorias “Utilidade” e “Custos”, o destaque recai sobre a folksonomia, com as maiores pontuações. Em ambas as categorias os demais sistemas de classificação possuem avaliação semelhante.

Com isto, observa-se que ao se tratar da estrutura e organização do sistema de classificação, a folksonomia não é a melhor opção. No entanto, caso o enfoque seja na facilidade de classifi-

cação e de utilização, como também, no custo financeiro e dispêndio de tempo, em relação aos demais sistemas aqui avaliados, a folksonomia leva grande vantagem.

A partir de uma análise particular entre a folksonomia e cada um dos SOC anteriormente descritos, foi possível observar:

**Folksonomia vs Universal:** são sistemas maioritariamente opostos. Assemelham-se apenas quando se trata da cobertura das áreas temáticas e da classificação em diferentes línguas. Enquanto o esquema universal está direccionado para uma organização mais lógica e melhor estruturada, a folksonomia aponta para a simplicidade e facilidade na utilização, sem qualquer preocupação com o vocabulário ou com a estrutura. Outra grande diferença entre eles é o custo financeiro e o dispêndio de tempo. Enquanto os esquemas universais são bastante caros e necessitam de alocação de algum tempo para a sua implementação e manutenção, os sistemas folksonómicos são simples de implementar e manter, reflectindo custos muito baixos e até, por vezes, gratuitos.

**Folksonomia vs Facetado:** em relação à estrutura e classes, os dois sistemas são semelhantes. No caso da terminologia, a única diferença está no controlo do vocabulário. Enquanto na folksonomia não há controlo algum, num sistema facetado, a existência de atributos pré-definidos faz com que se tenha um mínimo de controlo sobre os termos atribuídos a cada um. Por exemplo, numa faceta destinada à datas não serão atribuídos valores numéricos ou de texto. Quando se trata dos critérios direccionados à informalidade, estes sistemas tornam-se distintos. A folksonomia, pelas suas características colaborativa e informal, cria de forma voluntária, ou involuntária, comunidades ligadas por assuntos ou termos em comum, como também proporciona uma pesquisa serendipista, ou seja, permite descobertas de temas de interesse ao acaso. No entanto, o mesmo não ocorre no caso dos sistemas facetados.

Como nos demais sistemas formais de classificação, o facetado também é feito por um grupo restrito de pessoas, as quais definem um número limitado de atributos a serem imputados a um elemento. No caso da folksonomia, a classificação é feita por qualquer pessoa – seja ou não um especialista – e o número de atributos a serem imputados é ilimitado. Se por um lado a folksonomia é mais flexível do que um sistema facetado, por outro é menos estruturada e menos precisa. Outro diferencial entre eles é o factor custo – financeiro e de tempo despendido –, mais elevado no desenvolvimento de um sistema facetado.

**Folksonomia vs Tesouro:** ambos possibilitam a cobertura de praticamente todas as áreas temáticas, possuem uma estrutura que permite incorporar o melhor vocabulário para descrever o conhecimento, estão aptos a uma classificação multilingue e, a partir da perspectiva do utilizador, são relativamente fáceis de navegar. No entanto, diferem quando é exigida alguma estrutura ou categorização, pois estes não são itens prioritários da folksonomia.

Se por um lado o tesouro possui uma estrutura melhor organizada, com pouca ou nenhuma subjectividade, por outro é pouco flexível e hospitaleiro. Neste ponto, a folksonomia está em vantagem, pois é totalmente flexível e a qualquer momento está apta a acomodar novos temas. Do ponto de vista do classificador, o tesouro é uma ferramenta mais complexa e trabalhosa do que a folksonomia. Enquanto para manusear um tesouro é necessário ter algum conhecimento, a folksonomia é basicamente intuitiva, não exigindo um grande esforço intelectual. O factor custo é também um diferencial entre estes sistemas.

**Folksonomia vs Ontologia:** contrariamente à folksonomia, a ontologia é uma forma de organização onde todos os elementos são distintos e possuem conceitos e relacionamentos claros. Embora ambas possuam uma estrutura flexível e adaptável à incorporação de novos temas, para a folksonomia esta é uma tarefa fácil, feita com a simples adição de termos que referenciem o novo elemento. Para a ontologia o processo é mais complexo, sendo necessário conceptualizar e relacionar o novo tema de forma clara e única. Se na ontologia a estrutura, as classes e os relacionamentos estão bem organizados e não há subjectividade, na folksonomia isto não ocorre, pois não há controlo algum sobre a estrutura e sobre os termos utilizados. Na ontologia o controlo é feito por especialistas, enquanto que na folksonomia são os próprios utilizadores que definem como cada elemento será organizado, sendo este um dos seus factores de sucesso.

Em relação a descobertas ao acaso, classificação colaborativa e criação de comunidades através de assuntos de interesse – como foi observado também em relação aos demais sistemas de classificação –, a folksonomia é o único sistema que possui estas características. Relativamente ao factor custo, a folksonomia revela-se mais vantajosa do que a ontologia, na economia e rapidez de implementação.

**Folksonomia vs Taxonomia:** de acordo com a grelha de comparação aqui utilizada, a taxonomia possui quase as mesmas características da ontologia. Desta forma, pode-se dizer que as diferenças entre a folksonomia e a taxonomia são semelhantes às diferenças entre a folksonomia e a ontologia. A única distinção é o menor grau de complexidade da taxonomia em relação à ontologia, e consequentemente em relação à folksonomia.

Num contexto geral, a folksonomia destaca-se pela simplicidade e facilidade de uso. Mesmo sendo um sistema de classificação pouco estruturado e menos preciso, é rapidamente aceite pelos utilizadores por se adequar às necessidades de organização de cada um. Face ao grande volume de informações com que actualmente as pessoas se deparam, talvez seja mais agradável e eficaz utilizar uma classificação simples que satisfaça grande parte das necessidades, do que algo mais elaborado, que exija um maior consumo de tempo. Além disto, a folksonomia

não exige grandes conhecimentos para o seu manuseamento, ao contrário dos demais sistemas de classificação.

A descoberta ao acaso também pode ser um atractivo ao seu uso, pois torna a pesquisa, por vezes, uma tarefa um pouco mais agradável. Outro factor interessante da folksonomia é a sua interactividade com os utilizadores. Ao mesmo tempo que permite classificar e pesquisar informações, possibilita localizar pessoas com os mesmos assuntos de interesse. Do ponto de vista do utilizador, a falta de regras para a classificação pode ser um factor favorável, pois torna-se uma tarefa praticamente intuitiva.

Sobretudo para as organizações, um factor a considerar é o custo, quer financeiro quer de consumo de tempo, tornando-se, muitas vezes, um aspecto primordial na decisão do sistema de classificação a implementar. No caso dos sistemas formais, o custo pode, frequentemente, revelar-se bastante elevado, dificultando a sua adopção. Em face disto e pelo seu baixo investimento, a folksonomia pode ser considerada uma boa alternativa.

No entanto, convém referir que apesar dos aspectos positivos, em termos de estrutura, organização e recuperação de informações, a folksonomia não é o substituto para os sistemas formais de classificação, mas uma boa alternativa quando a pesquisa se restringe a algo simples e colaborativo. A complementaridade entre um sistema formal e a folksonomia poderia resultar numa boa opção.

## **6 COMPORTAMENTO DE PESQUISA DE INFORMAÇÃO DE ENGENHEIROS E CIENTISTAS**

Para o desenvolvimento do presente estudo, além das informações abordadas nos capítulos anteriores relacionadas com as características, vantagens e desvantagens dos SOC, é importante também compreender algumas particularidades do ambiente no qual ele, ou eles, poderão ser implementados.

O estudo pode ser realizado tanto de uma perspectiva mais genérica, como no estudo de (Damme 2006) – chamado “Folksonomies and Enterprise Folksonomies” – onde a autora avaliou os tipos de organização, ou através de uma análise mais específica, estudando, por exemplo, um ambiente particular, como neste caso, um instituto de I&D.

A motivação para avaliar um instituto de I&D, a partir da perspectiva do comportamento de pesquisa de informação de engenheiros e cientistas explica-se uma vez que se tomou em consideração dois factores principais: a) neste trabalho, o que está em questão são SOC formais e informais, ou seja, alguns deles (informais) são desenvolvidos e mantidos pelos próprios utilizadores, e, neste caso, é importante conhecer o comportamento dos mesmos na pesquisa de informação; b) o tipo de profissionais que actuam no instituto apresentado na secção 6.2, identifica-se, de uma forma geral, com engenheiros e cientistas.

### ***6.1 Estudos Anteriores***

O conceito de “pesquisa de informação” é entendido por (Hertzum 2008) como actividades envolvidas na percepção de uma necessidade de informação, ou seja, a exploração desta necessidade, a identificação, selecção e consulta de fontes de informação, adaptação da informação aos problemas do trabalho, como também toda a informação proveniente sem activamente estar à procura dela. Estas actividades ocorrem num contexto cooperativo mas podem ser realizadas por intervenientes individuais ou diversos intervenientes em cooperação.

Para (Wilson 2000), o comportamento de pesquisa de informação resulta de uma necessidade de satisfação de um objectivo. No decorrer da pesquisa, a interacção pode ser feita utilizando sistemas de informação manuais (tais como publicações numa biblioteca) ou através de sistemas informáticos (tais como a Internet).

Os primeiros estudos sobre o comportamento de pesquisa de informação podem ser encontrados em trabalhos realizados com utilizadores de bibliotecas e estudos gerais sobre leitores.

O ano de 1948 foi considerado como o início do estudo moderno sobre o comportamento de pesquisa de informação, com a Royal Society Scientific Information Conference. No entanto, as primeiras referências encontradas sobre o tema são anteriores a 1916 (Wilson 2000).

Ao longo de várias décadas têm sido conduzidos estudos sobre o comportamento de pesquisa de informação dos engenheiros e cientistas. Os pesquisadores têm desenvolvido diferentes tipos de modelos, estando alguns focados na identificação específica das actividades do comportamento de pesquisa, e outros numa visão mais holística, através da incorporação de factores que influenciam no processo da pesquisa de informação. Adicionalmente, alguns modelos exploram o comportamento de pesquisa de informação em diferentes contextos (Hirsh and Dinkelacker 2004).

Seguindo esta metodologia de abordagem, serão apresentados alguns trabalhos realizados sobre o comportamento de pesquisa de cientistas e engenheiros e alguns dos seus principais resultados.

Em 1997, (Ellis and Haugan 1997) realizaram um estudo sobre a natureza da comunicação em trabalhos de pesquisa e regras de comportamento de pesquisa relacionadas com as tarefas de trabalho envolvidas no projecto. Neste trabalho, os autores distinguiram o comportamento dos cientistas e dos engenheiros.

No caso dos cientistas, puderam observar que:

- O comportamento de pesquisa é mais intenso no início do projecto, onde ambos os canais – formais e informais – são utilizados;
- A utilização dos canais formais diminui de acordo com o avanço do projecto, e a comunicação interpessoal parece ser um factor predominante;
- No final do projecto ambos os canais – formais e informais – são reutilizados, mas em menor escala;
- Na maioria das vezes, os cientistas procuravam informações externas e, comparado aos engenheiros, utilizavam com maior frequência o contacto com outros estudiosos para se manterem actualizados;
- Para assuntos não familiares, os cientistas escolhiam como ponto de partida pesquisas em aplicações informáticas por literatura retrospectiva, seguido por contactos pessoais noutros ambientes científicos;

- Os cientistas utilizam mais frequentemente serviços de informação online do que os engenheiros.

Em relação aos engenheiros, destacam-se os seguintes resultados:

- Utilizam mais a comunicação dentro do seu próprio departamento ou equipa de projecto, ou a própria organização como um todo, para se manterem actualizados;
- A documentação interna foi identificada como um dos mais importantes recursos;
- Para assuntos não familiares, a opção seria a pesquisa entre colegas ou outras pessoas dentro da sua rede de contactos, seguido do apoio de um bibliotecário na pesquisa em aplicações informáticas por literatura retrospectiva;
- Os canais de informação preferenciais eram baseados na sua experiência e conhecimento;
- Em relação às tarefas técnicas, os engenheiros utilizavam com maior frequência revistas e relatórios internos do que publicações científicas e conferências.

Ainda sobre estes dois intervenientes no estudo, cientistas e engenheiros, (Anderson, Glassman et al. 2001) realizaram um estudo no sector aeroespacial dos Estados Unidos. Apesar de não diferenciar a preferência de cada uma das categorias de profissionais estudadas, os resultados assemelham-se aos obtidos por (Ellis and Haugan 1997).

Na escolha do suporte de informação, os autores observaram que:

- A preferência ficou pela colecção pessoal e comunicação oral dentro da organização, seguida pela comunicação com pessoas externas à organização;
- As escolhas menos utilizadas foram a literatura e um técnico de biblioteca;
- Com a crescente incerteza nas actividades a realizar, a preferência pela comunicação oral diminui, aumentando a opção pela pesquisa literária e pelo recurso ao técnico de biblioteca, sendo que não foram estes os factores que mais influenciaram na selecção de suportes impressos. O mesmo se poderá concluir em relação à acessibilidade e qualidade;

- Assim sendo, resulta como primeiro determinante na utilização de suportes impressos a importância deste material para o trabalho do utilizador.

Num outro estudo, (Hertzum and Pejtersen 2000) aborda o comportamento de pesquisa dos engenheiros na busca de informação científica, em manuais e relatórios e através da comunicação interpessoal. Para os autores, os engenheiros pesquisam documentos para encontrar pessoas, procuram pessoas para encontrar documentos, e interagem socialmente para conseguir informação oral e escrita, sem participarem directamente em pesquisas. No entanto, para obter esta informação deparam-se com um obstáculo, o custo/tempo.

Nas áreas de design, processos e produção, (Kwasitsu 2003) desenvolveu um estudo com engenheiros de uma companhia internacional de fabrico de micro-chips. Em relação ao comportamento de pesquisa de informação, o autor pôde observar que:

- A solução de problemas é a actividade que mais desencadeia a pesquisa de informação, seguida por investigação de uma nova ideia e planeamento de um projecto;
- Em relação à importância dos vários recursos de informação, a maioria dos engenheiros considerou as pessoas dentro do seu próprio grupo de trabalho como recursos de informação altamente importantes, seguida dos aspectos cognitivos, colecção pessoal e Internet;
- A acessibilidade e disponibilidade foram características indicadas como altamente importantes na influência da selecção de recursos de informação;
- Saber onde procurar foi o problema mais referido no processo de aquisição de informação, seguido por excesso de informação e informação indisponível.

Em relação à preferência por recursos de informação, (Hertzum 2002) elaborou um estudo sob o ponto de vista da importância da confiança dos engenheiros de software na avaliação e escolha desses elementos. Para o autor, a confiança é de suma importância, pois a qualidade é uma propriedade percebida e, assim, avaliar a qualidade de um recurso de informação é essencialmente uma questão de estabelecer até que ponto está disposto a confiar nele.

Isto pode ser feito com grande facilidade e precisão utilizando recursos familiares. O conceito de confiança, segundo o autor, fornece uma explicação alternativa pela preferência dos engenheiros por recursos de informação internos.

Este estudo permitiu a (Hertzum 2002) perceber que:



- No estudo de projectos de desenvolvimento de sistemas, os engenheiros de software podem avaliar com mais confiança a idoneidade das informações provenientes dos recursos que lhe são familiares;
- Inicialmente procuravam informações em fontes que eram facilmente acessíveis, ou seja, recursos internos e recursos que já eram conhecidos;
- Sistemas acessíveis em qualquer lugar e a qualquer hora, o que se traduz em custos reduzidos, não são suficientes para criar uma relação de familiaridade entre os engenheiros e o próprio sistema. Além disto, é necessário fornecer informações que permitam que os engenheiros avaliem a qualidade das informações disponíveis;
- Os factores que determinam a escolha de colegas como fontes de informação são: acessibilidade, familiaridade e facilidade para determinar a sua veracidade;
- Os engenheiros de software procuram informações acessíveis, de uma forma que permita perceber a sua qualidade;
- A atenção recai sobre o factor qualidade do que sobre o factor custo da fonte de informação.

Em relação ao factor acessibilidade, (Fidel and Green 2004) analisaram as suas diferentes componentes na percepção dos engenheiros sobre os recursos de informação. Para eles, nos estudos anteriores sobre o comportamento de pesquisa de informação dos engenheiros, a acessibilidade era o factor que mais influenciava na selecção de recursos. No entanto, segundo a sua perspectiva, o conceito de acessibilidade é ambíguo, tendo dado lugar a diferentes interpretações pelos cientistas e engenheiros.

Foi utilizando os relatos dos engenheiros, que (Fidel and Green 2004) procuraram identificar os factores que mais influenciavam na escolha de um recurso de informação.

Com base no estudo realizado puderam concluir:

- Ao contrário de estudos anteriormente realizados, a proximidade física não foi a principal razão para considerar colegas de trabalho como os principais recursos de informação. Segundo os autores, este conceito deu lugar ao factor familiaridade, pois é este que contribui para minimizar o esforço intelectual;
- Assim sendo, a familiaridade interfere activamente no factor de selecção dos colegas de trabalho como fontes de informação, sendo relevante estabelecer laços

entre estes dois intervenientes. Em contraste, relativamente às fontes de informação impressas, a principal motivação visa a optimização de tempo.

Importa referir que, no âmbito do estudo do comportamento de pesquisa de informação, (Hirsh and Dinkelacker 2004) e (Guo and Li 2007) elaboraram trabalhos centrados na temática Investigação e Desenvolvimento. Um dos estudos realizados envolveu especialistas de um instituto de pesquisa e o outro, profissionais de I&D na criação de novos produtos.

Segundo (Hirsh and Dinkelacker 2004), a principal característica de quem trabalha com o conhecimento é a utilização da informação para gerar mais informação. Assim, a pesquisa de informação é um aspecto central do seu quotidiano profissional. Em institutos de pesquisa esta tendência torna-se mais evidente, pelo facto dos resultados produzidos serem apresentados como informação, sob a forma de publicações, relatórios técnicos, patentes, protótipos etc.

No decorrer do estudo elaborado com os profissionais da Hewlett Packard e da Compaq Computer, os autores procuraram identificar a forma de utilização das fontes de informação, as preferências por estes recursos e os tipos de fontes de informação por eles produzidos.

Com base neste estudo, os autores identificaram os seguintes resultados:

- A utilização da Internet foi ferramenta predominante na pesquisa de informações;
- Na resolução de problemas, ao contrário de outros estudos, onde engenheiros e cientistas consideraram colegas de trabalho como fontes de informação e as primeiras referências, os autores verificaram que, neste caso, a preferência recaiu sobre indivíduos fora da instituição;
- O tempo dispendido na monitorização das informações e a autenticidade dos recursos foram os factores que mais influenciaram na selecção de fontes de informação.

Ainda no âmbito de I&D, (Guo and Li 2007) realizaram um estudo sobre o comportamento de pesquisa de informação dos profissionais ligados a esta área, no desenvolvimento de novos produtos. Neste artigo, os autores examinaram em pormenor de que forma é que as características individuais, de tarefas e das fontes de informação afectaram o comportamento de pesquisa de informação desses profissionais. No final do estudo, concluíram que:

- O acréscimo de incertezas e a complexidade de tarefas aumenta a utilização dos múltiplos recursos de informação;

- As fontes internas, impressas e os recursos de rede são factores significativamente relacionados com a qualidade e acessibilidade da informação;
- Relativamente às características individuais, a formação académica e a experiência profissional estão associadas à frequência e à dimensão da utilização da fonte de informação;
- Os profissionais de I&D de diferentes campos técnicos ou mesmo de diferentes áreas funcionais, são solicitados a realizar uma variedade de actividades em paralelo, interagindo uns com os outros e frequentemente partilhando informações.

O comportamento de pesquisa de informação tem sido analisado durante décadas. No caso dos engenheiros e cientistas, (Anderson, Glassman et al. 2001) apresentaram algumas conclusões gerais, resultado de diversos trabalhos realizados até então. Para eles, existe uma ligação entre a performance da empresa e a eficácia da sua comunicação. Além disso, a comunicação técnica interna é uma importante fonte de inovação. Para reduzir as incertezas e complexidades, as organizações utilizam informações técnicas e de negócios, na maioria das vezes provenientes de ambientes externos.

Concluíram por fim que existe uma preferência pela comunicação interpessoal quando perante informações não rotineiras, complexas e ambíguas, e uma preferência por suportes impressos quando perante mensagens de rotina, claras e simples.

## ***6.2 O comportamento na selecção, organização, classificação e partilha de informação: um estudo sobre o INESC-Porto***

Não é recente o estudo sobre comportamento de pesquisa de informação de profissionais da área de I&D, seja em ambientes empresariais ou educacionais. Pelo exposto no capítulo 6, este tema foi já objecto de vários trabalhos. No entanto, o processo de selecção, organização, classificação e partilha de informação, conjugando meios formais, informais e também sistemas que possibilitam desenvolver estes tipos de processo de forma simples, intuitiva e colaborativa, como a folksonomia, por exemplo, é ainda um campo de estudo pouco explorado.

O objecto deste estudo versa, pois, sobre o tipo de comportamento anteriormente abordado. Tendo como base os trabalhos apresentados neste capítulo e na análise de algumas características dos SOC, procedeu-se à elaboração de um questionário online. O desenho do mesmo foi elaborado recorrendo a uma ferramenta específica *Web* para a elaboração de inquéritos,

permitindo a recolha de dados e análise dos seus resultados. O seu endereço de acesso foi enviado por e-mail aos inquiridos, tendo como base uma estimativa no tempo de resposta entre 10 e 15 minutos para cada questionário e uma dilação de 60 dias para a recolha dos dados.

A adopção de um inquérito online com perguntas fechadas teve como objectivo abranger o maior número de indivíduos e assegurar uma interpretação fácil e intuitiva do seu conteúdo, garantindo, assim, uma resposta rápida e fiável. Nele, foram introduzidos três grupos de questões: a) informações profissionais e académicas; b) organização e partilha de informação; c) fontes de informação.

### **6.2.1 Objectivos**

O objectivo deste estudo foi perceber como os profissionais na área de I&D, que actuam directa<sup>10</sup> ou indirectamente<sup>11</sup> no INESC-Porto, seleccionam, organizam, classificam e partilham as suas informações.

Mais especificamente, no seguimento do que foi delineado, procurou-se identificar:

- Qual o conhecimento que esses profissionais têm sobre sistemas formais e informais de organização do conhecimento.
- Como organizam, classificam, guardam e partilham, por via electrónica, documentos e conteúdos.
- Quais as fontes de informação por eles utilizadas, com que frequência as utilizam para o seu trabalho e quais são os factores que mais influenciam na sua escolha e acesso.
- As principais dificuldades encontradas na pesquisa por informação.

### **6.2.2 Público-alvo**

O estudo foi dirigido aos responsáveis de projecto, investigadores, bolsiros e todos os demais colaboradores do INESC-Porto no exercício da sua actividade técnico-científica.

---

<sup>10</sup> Indivíduos vinculados a algum departamento do INESC-Porto.

<sup>11</sup> Indivíduos sem vínculo aos departamentos do INESC-Porto.

### 6.2.2.1 INESC-Porto

O INESC-Porto (INESC-Porto 2008) é uma associação privada sem fins lucrativos, reconhecida como instituição de utilidade pública, tendo adquirido em 2002 o estatuto de Laboratório Associado.

Desenvolve actividades de I&D, consultoria, formação avançada e transferência de tecnologia nas áreas de Telecomunicações e Multimédia, Sistemas de Energia, Sistemas de Produção, Sistemas de Informação e Comunicação e Optoelectrónica.

O INESC-Porto LA (Laboratório Associado) possui ainda dois grupos autónomos, o Laboratório de Inteligência Artificial e Apoio à Decisão (LIAAD) e o Centro de Investigação em Sistemas Computacionais Avançados (CRACS).

### 6.2.2.2 Inquiridos

O presente inquérito foi distribuído por um universo de 228 indivíduos, directa ou indirectamente ligados ao INESC-Porto. Registaram-se 36 (15,7%) respostas, sendo que destas, 30 (83,3%) correspondem ao preenchimento do inquérito na sua totalidade.

A tabela elaborada e apresentada em baixo reflecte o universo de respostas obtidas, com uma breve caracterização dos indivíduos. Para assegurar a confidencialidade das respostas, definiu-se um número para cada indivíduo e o nome do departamento ao qual pertence foi substituído por uma letra. O indicador para aqueles que não estão directamente ligados a um departamento do INESC-Porto é “Externo”.

**TABELA 2 – Informações sobre os inquiridos**

Inquiridos	Formação	Departamento	Função	Experiência (anos)
1	Mestrado	A	Investigador	2
2	Doutoramento	A	Investigador	1,6
3	Mestrado	B	Investigador / docente	6
4	Licenciatura	B	Consultor	2
5	Mestrado	B	Investigador	7
6	Mestrado	B	Investigador	7
7	Mestrado	C	Coordenador adjunto	8
8	Licenciatura	A	Bolseiro	1

Continua...

...Continuação

<b>9</b>	Mestrado	B	Colaborador	< 1
<b>10</b>	Licenciatura	C	Investigador / aluno de mestrado	1
<b>11</b>	Mestrado	Externo	Bolseiro	< 1
<b>12</b>	Mestrado	C	Bolseiro de doutoramento	2
<b>13</b>	Doutoramento	A	Docente	9
<b>14</b>	Doutoramento	D	Docente	8
<b>15</b>	Mestrado	B	Investigador	1
<b>16</b>	Mestrado	C	Project Engineer	5
<b>17</b>	Mestrado	C	Bolseiro	2
<b>18</b>	Licenciatura	B	Bolseiro programador	< 1
<b>19</b>	Mestrado	G	Engenheiro	4
<b>20</b>	Mestrado	C	Bolseiro de investigação	1
<b>21</b>	Licenciatura	A	Estagiário	< 1
<b>22</b>	Licenciatura	C	Bolseiro de investigação	2
<b>23</b>	Mestrado	C	Bolseiro de investigação	1
<b>24</b>	Licenciatura	B	Responsável de projecto	13
<b>25</b>	Mestrado	B	Investigador	3
<b>26</b>	Doutoramento	C	Líder de área	15
<b>27</b>	Licenciatura	B	Responsável de projecto	12
<b>28</b>	Doutoramento	B	Investigador	11
<b>29</b>	Licenciatura	B	Analista / programador	1
<b>30</b>	Mestrado	E	Investigador convidado	1
<b>31</b>	Licenciatura	E	Coordenação	2
<b>32</b>	Doutoramento	F	Docente	25
<b>33</b>	Doutoramento	D	Professor	30
<b>34</b>	Mestrado	Externo	Docente	10
<b>35</b>	Mestrado	B	Investigador	7
<b>36</b>	Doutoramento	G	Investigador sénior	7

Do universo de 36 indivíduos que responderam ao inquérito, ressalta diferente experiência profissional e formação académica. A grande maioria (18 pessoas, ou 50%) possuía mestrado, seguido por licenciatura (10 pessoas, ou 27,8%) e doutoramento (8 pessoas, ou 22,2%). Outro aspecto relevante verificado foi a variação do tempo de experiência profissional dos intervenientes. A média total foi de 8,5 anos no exercício profissional, com variação entre <1 e 30 anos. Apenas 7 indivíduos (19,4%) ultrapassaram os 10 anos de experiência.

Do total de 36 indivíduos, apenas 5,5% (ou 2 pessoas) eram externos ao INESC-Porto. Os demais estavam vinculados a sete departamentos diferentes. O departamento “B” foi o que obteve a maior participação no inquérito, com 36,1% (ou 13 indivíduos), seguido por “C”

(25%, ou 9 indivíduos), “A” (13,8%, ou 5 indivíduos), “D” (5,6%, ou 2 indivíduos), “E” (5,6%, ou 2 indivíduos), “G” (5,6%, ou 2 indivíduos) e “F” (2,8%, ou 1 indivíduo).

### **6.2.3 Resultados**

Dos três grupos de questões analisados neste inquérito, o primeiro foi já discutido anteriormente no ponto 6.2.2. Nesta secção será, pois, apresentada uma análise dos dois últimos grupos de questões, que versaram sobre a temática da organização e partilha de informação e das fontes de informação.

#### **6.2.3.1 Organização e Partilha de Informação**

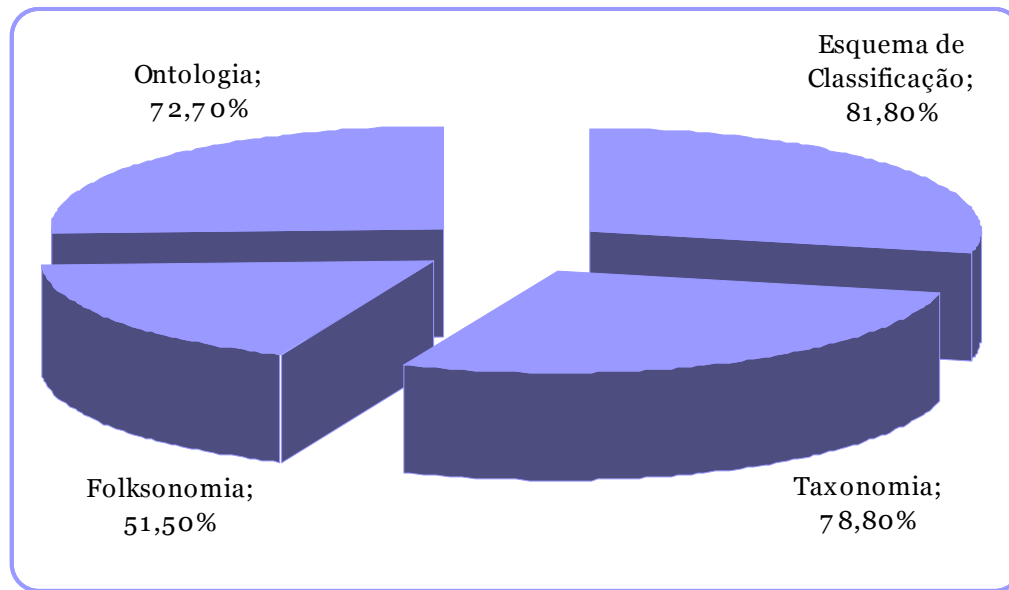
Tendo como objectivo analisar o processo de organização e partilha de informação, foram elaboradas três questões mais específicas. Com a primeira questão pretendia-se aferir do conhecimento dos inquiridos sobre alguns sistemas – formais e informais – de organização do conhecimento. Em seguida procurou-se identificar os métodos de organização, classificação e armazenamento de documentos e conteúdos. A última questão versou sobre a partilha, por via electrónica, de documentos e conteúdos.

Com os resultados obtidos nestas questões, foi possível elaborar a seguinte análise:

**Questão 1:** Indique se conhece ou já ouviu falar sobre esquema de classificação, taxonomia, folksonomia ou ontologia (para cada opção foi apresentado um breve conceito).

Na sua grande maioria pode constatar-se que havia já um conhecimento mais ou menos empírico do esquema de classificação (81,8%, ou 27 pessoas), seguida pela taxonomia (78,8%, ou 26 pessoas) e pela ontologia (72,7%, ou 24 pessoas). A folksonomia – dentre todas as opções, o único sistema informal de organização do conhecimento – constatou-se ser a opção menos familiar (51,4%, ou 17 pessoas).

**GRÁFICO 1 – Conhecimento sobre sistemas de organização do conhecimento**



Com base no estudo elaborado, que consubstancia este trabalho, uma das relações que se pode retirar é que a folksonomia, por se tratar de um tema recente, ainda pouco explorado na literatura científica e uma ferramenta raramente aplicada dentro de empresas e instituições, apesar da sua cada vez maior utilização na *Internet*, é naturalmente o SOC menos popular entre os anteriormente citados.

Esta lacuna no conhecimento teórico da folksonomia, não significa, no entanto, que não tenha sido aplicada, mesmo que involuntariamente. A proliferação deste conceito em sites da *Web* tem sido uma constante. Além dos *sites* como Delicious e Flickr, já mencionados em capítulos anteriores, vários outros utilizam sistemas baseados em *tags*. O Gmail – servidor de e-mails da Google –, por exemplo, não organiza as mensagens numa estrutura hierárquica ou em pastas, mas possui um motor de pesquisa próprio, como também permite que os utilizadores atribuam a sua terminologia pessoal (*tags*) para classificar cada mensagem.

Através deste processo, é possível pesquisar tanto pelos elementos da mensagem (remetentes, destinatário(s), assunto, corpo do texto etc.), quanto pelas *tags* atribuídas pelos utilizadores. Esta atribuição de *tags* cria automaticamente uma estrutura flexível e personalizada (não hierárquica) de assuntos, substituindo a hierarquia em pastas, por sua vez menos flexível do que uma classificação mais informal.

**Questão 2:** Quer individualmente, quer colectivamente, como organiza/classifica/guarda documentos e conteúdos?



Das 35 respostas obtidas a esta, questão, concluiu-se que os indivíduos de alguma forma organizam, classificam e guardam os seus conteúdos. Destaca-se a utilização de pastas do próprio computador como elemento preferencial de 97,1% (ou 34 pessoas), seguida pela utilização de *bookmarking* (23 pessoas, ou 65,7%). Numa posição intermédia encontrou-se a utilização de aplicações próprias para a gestão de referências bibliográficas e documentos electrónicos (37,1%, ou 13 pessoas), e utilização de wikis (25,7%, ou 9 pessoas). Ferramentas como *social bookmarking* (11,4%, ou 4 pessoas), funcionalidades do gestor de blog (5,7%, ou 2 pessoas), discos virtuais (5,7%, ou 2 pessoas) e ferramentas de apoio à classificação (5,7%, ou 2 pessoas) foram menos citadas pelos inquiridos. Além das opções pré-definidas, uma das respostas referia os ficheiros de texto (txt) como uma forma de organização de informação.

**TABELA 3 – Organização, classificação e guarda de documentos e conteúdos**

<b>Opções</b>	<b>n</b>	<b>%</b>
Não organizo nem classifico documentos e conteúdos por falta de tempo	0	0.0%
Não organizo nem classifico documentos e conteúdos porque não necessito	0	0.0%
Uso pastas no próprio computador para organizar, classificar e guardar documentos e conteúdos (ex.: por assuntos, ordem alfabética, numérica, cronológica etc.)	34	97.1%
Uso a intranet (ou semelhante) da minha organização e as suas funcionalidades de organização e classificação	7	20.0%
Uso ferramentas de apoio à classificação dos documentos (ex.: thesaurus, cabeçalhos de assunto, esquemas de classificação etc.)	2	5.7%
Não me preocupo muito com isso porque uso um motor de pesquisa no meu computador pessoal ou intranet (ex.: Google Desktop, Windows Sidebar, Spotlight, sistema de pesquisa da própria plataforma, etc.)	4	11.4%
Uso uma aplicação própria para a gestão de referências bibliográficas e documentos electrónicos (tipo EndNote, JabRef, Papers)	13	37.1%
Uso Bookmarking (ex.: favoritos do Internet Explorer, Mozilla Firefox, favoritos da barra de ferramentas do Google, barra de ferramentas do Yahoo etc.)	23	65.7%
Social bookmarking (ex.: CiteUlike, del.icio.us, Flickr, Furl, ma.gnolia, Technorati, Yahoo My Web etc.)	4	11.4%
Uso Wikis (ex.: TWiki, TikiWiki, Wikia, PBWiki, Wikispaces, JotSpot etc.)	9	25.7%
Uso as funcionalidades do gestor do meu Blog (ex: Typepad, Blogspot)	2	5.7%
Uso discos virtuais (ex.: Sky Drive da Microsoft etc.)	2	5.7%
Uso ferramentas de produtividade colaborativas (ex.: Google Docs & Spreadsheets, Zoho Writer, ajaxWrite etc.)	6	17.1%

Outro factor observado foi que, apesar de 51,5% (ou 17 pessoas) dos inquiridos conhecerem ou já terem ouvido falar da folksonomia, apenas quatro a utilizavam na prática, correspondendo a 23,5% deste total. Em relação aos esquemas de classificação, a diferença registada foi maior, já que (81,8%, ou 27 pessoas) conhecia ou já tinha ouvido falar, mas apenas 7,4% (ou

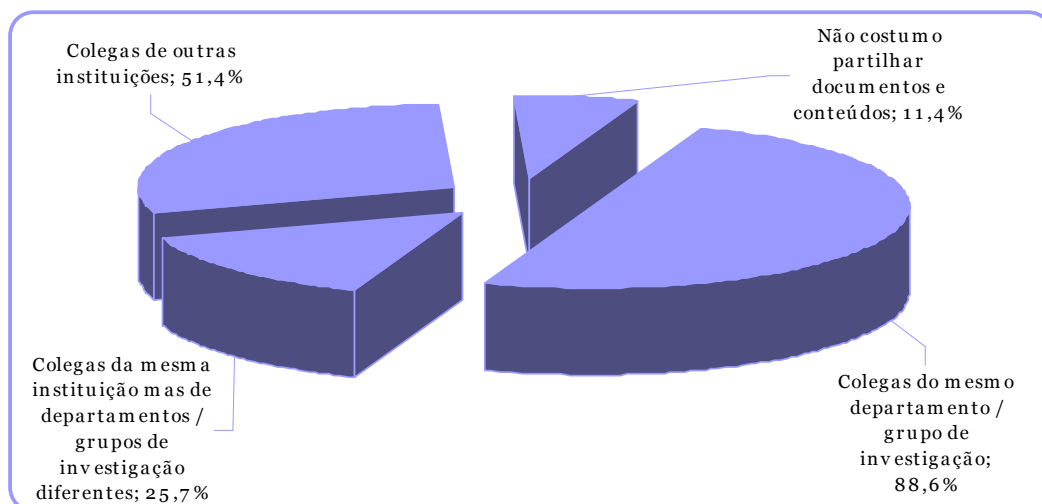
2 pessoas) os utilizavam efectivamente. A usabilidade, familiaridade e tempo dispendido podem ser factores possivelmente associados a tal facto, pois, como mencionado em capítulos anteriores, apesar de menos precisa, a folksonomia é uma maneira mais rápida, simples e fácil de organizar conteúdos do que os tradicionais esquemas de classificação.

A utilização de *bookmarking* por grande parte dos inquiridos foi um indicador de que a *Internet* é uma tecnologia bastante familiar e habitualmente utilizada. Contudo, outras ferramentas associadas à *Web*, que apoiam a partilha de informação e o trabalho colaborativo, como *social bookmarking*, *wikis* e ferramentas de produtividade colaborativa (ex.: Google Docs), são ainda pouco exploradas.

**Questão 3:** Com quem costuma partilhar, por via electrónica, documentos e conteúdos?

Apenas 35 inquiridos responderam a esta questão. A grande preferência recaiu sobre colegas do mesmo departamento/grupo de investigação, com 88,6% (ou 31 pessoas). Em segundo lugar foram mencionados os colegas de outras instituições (51,4%, ou 18 pessoas), superando a opção colegas da mesma instituição mas de departamentos/grupos de investigação diferentes, que obteve 25,7% (ou 9 pessoas) da preferência dos inquiridos. Apesar da investigação ser uma área onde a informação é a matéria-prima principal, alguns inquiridos (11,4%, ou 4 pessoas) indicaram não partilhar usualmente documentos e conteúdos.

**GRÁFICO 2 – Pessoas com quem partilha, por via electrónica, documentos e conteúdos**



Da leitura do gráfico 2, pudemos concluir que a instituição não era o principal elo de ligação entre os inquiridos. Exceptuando os colegas do mesmo grupo de investigação, onde a troca de

informação era quase que uma regra, a preferência foi pela partilha de informação com pessoas externas à instituição. Este não é um facto invulgar em I&D, pois temas de interesse são factores que unem pessoas ligadas a uma mesma área, mas que nem sempre trabalham na mesma instituição.

### **6.2.3.2 Fontes de Informação**

O último bloco de questões colocadas, referente às fontes de informação, pretendeu esclarecer igualmente algumas dúvidas. Na primeira questão o objectivo era identificar quais e com que frequência, eram as fontes de informação utilizadas para obter o conteúdo necessário ao desenvolvimento do trabalho de investigação. Em seguida, procurou-se esclarecer o grau de influência de alguns factores na escolha e acesso a estas fontes. Por fim, inquiriu-se sobre as dificuldades encontradas na pesquisa de informação para o desenvolvimento do trabalho de investigação.

**Questão 1:** Com que frequência utiliza algumas fontes – indicadas no inquérito – para obter a informação necessária para o desenvolvimento do trabalho de investigação?

De um total de 36 inquiridos, apenas 30 assinalaram esta questão. Para a obtenção de informação necessária ao trabalho de investigação, metade deles (15 pessoas, ou 50%) indicou recorrer com frequência à pesquisa na *Internet* de uma forma não organizada, seguida por publicações externas (46,7%, ou 14 pessoas), como artigos, periódicos científicos etc. Colegas do mesmo departamento/grupo de investigação (46,7%, ou 14 pessoas), biblioteca pessoal (43,3%, ou 13 pessoas) e outros locais específicos na Internet (36,7%, ou 11 pessoas), como notícias online, grupos de discussão, blogs etc., destacaram-se entre as fontes frequentemente utilizadas.

Em relação às fontes de informação por vezes utilizadas, de referir os colegas externos à instituição (46,7%, ou 14 pessoas), colegas de outro departamento/grupo de investigação (46,7%, ou 14 pessoas) e biblioteca da instituição ou intranet (43,3%, ou 13 pessoas). Das fontes quase nunca, ou nunca utilizadas, ressaltaram as bases de dados electrónicas (33,3%, ou 10 pessoas) e biblioteca externa à instituição (30%, ou 9 pessoas). Além das opções apresentadas no inquérito, um dos inquiridos assinalou a fonte de informação Scopus.

**TABELA 4 – Frequência de utilização das fontes de informação**

Fontes de Informação	Sempre	Frequentemente	As vezes	Quase nunca	Nunca	Respostas
Colegas do mesmo departamento / grupo de investigação	16,7% (5)	<b>46,7%</b> <b>(14)</b>	30,0% (9)	6,7% (2)	0,0% (0)	30
Colegas de outros departamentos / grupos de investigação	6,7% (2)	3,3% (1)	<b>46,7%</b> <b>(14)</b>	33,3% (10)	10,0% (3)	30
Colegas externos à instituição	6,7% (2)	16,7% (5)	<b>46,7%</b> <b>(14)</b>	26,7% (8)	3,3% (1)	30
Biblioteca pessoal (ex.: artigos, livros e outras publicações profissionais)	30,0% (9)	<b>43,3%</b> <b>(13)</b>	26,7% (8)	0,0% (0)	0,0% (0)	30
Biblioteca da instituição ou intranet (excepto a biblioteca pessoal)	3,3% (1)	33,3% (10)	<b>43,3%</b> <b>(13)</b>	10,0% (3)	10,0% (3)	30
Biblioteca externa à instituição	3,3% (1)	26,7% (8)	<b>30,0%</b> <b>(9)</b>	<b>30,0%</b> <b>(9)</b>	10,0% (3)	30
Fontes na internet pesquisadas de uma forma organizada (ex: MetaLib, B-on, etc.)	23,3% (7)	<b>30,0%</b> <b>(9)</b>	23,3% (7)	13,3% (4)	10,0% (3)	30
Fontes na internet pesquisadas de uma forma não organizada (ex: Google, Google Scholar, etc.)	<b>50,0%</b> <b>(15)</b>	46,7% (14)	3,3% (1)	0,0% (0)	0,0% (0)	30
Outros locais específicos na internet (ex.: notícias online, grupos de discussão, blogs, sites de projectos etc.)	10,0% (3)	<b>36,7%</b> <b>(11)</b>	23,3% (7)	23,3% (7)	6,7% (2)	30
Publicações externas (ex.: artigos, periódicos científicos etc.)	<b>46,7%</b> <b>(14)</b>	30,0% (9)	16,7% (5)	6,7% (2)	0,0% (0)	30
Bases de dados electrónicas específicas (ex.: Compendex, Inspec, etc.)	3,3% (1)	13,3% (4)	30,0% (9)	20,0% (6)	<b>33,3%</b> <b>(10)</b>	30

Na obtenção de informação para o desenvolvimento do trabalho de investigação, além dos colegas do mesmo grupo de trabalho, a *Internet* foi apontada como uma fonte frequentemente utilizada pelos inquiridos, principalmente através de motores de pesquisa, tais como Google, Yahoo!, Google Scholar etc. Publicações externas, como artigos e periódicos científicos, puderam também ser considerados, até certo ponto, como parte destas pesquisas numa forma não organizada, pois através dos próprios motores de pesquisa é possível recuperar diversos artigos científicos sem ter que recorrer a qualquer forma de índice ou catálogo de publicações.

Neste estudo, contrariamente ao trabalho realizado por (Hirsh and Dinkelacker 2004), colegas externos à instituição não foram citados como escolha frequente como fontes de informação. Na relação interpessoal, os inquiridos indicaram apenas colegas do mesmo departamen-

to/grupo de investigação como pessoas frequentemente consultadas na pesquisa e obtenção de informação.

**Questão 2:** Qual o grau de influência de alguns factores – indicados no inquérito – na escolha e acesso às fontes de informação electrónicas?

Esta questão obteve 30 respostas. Ao considerarmos os graus “muito importante” e “importante”, a boa qualidade técnica (100%) e a actualização (100%) foram os factores que mais influenciaram na escolha e acesso a uma fonte de informação electrónica, seguido por relevância (96,7%, ou 29 pessoas), fiabilidade (96,7%) e disponibilidade (96,7%).

**TABELA 5 – Factores que influenciam na escolha e acesso às fontes de informação electrónicas**

Factores	Muito importante	Importante	Pouco importante	Indiferente	Sem importância	Respostas
Boa qualidade técnica	<b>63,3%</b> <b>(19)</b>	36,7% (11)	0,0% (0)	0,0% (0)	0,0% (0)	30
Adequação às tarefas	40,0% (12)	<b>50,0%</b> <b>(15)</b>	6,7% (2)	3,3% (1)	0,0% (0)	30
Actualização	<b>63,3%</b> <b>(19)</b>	36,7% (11)	0,0% (0)	0,0% (0)	0,0% (0)	30
Relevância	<b>56,7%</b> <b>(17)</b>	40,0% (12)	3,3% (1)	0,0% (0)	0,0% (0)	30
Fiabilidade	<b>56,7%</b> <b>(17)</b>	40,0% (12)	3,3% (1)	0,0% (0)	0,0% (0)	30
Qualidade do motor de pesquisa	30,0% (9)	<b>40,0%</b> <b>(12)</b>	26,7% (8)	3,3% (1)	0,0% (0)	30
Qualidade na classificação dos assuntos	23,3% (7)	<b>60,0%</b> <b>(18)</b>	13,3% (4)	3,3% (1)	0,0% (0)	30
Acessibilidade	16,7% (5)	<b>73,3%</b> <b>(22)</b>	3,3% (1)	6,7% (2)	0,0% (0)	30
Disponibilidade	30,0% (9)	<b>66,7%</b> <b>(20)</b>	3,3% (1)	0,0% (0)	0,0% (0)	30
Usabilidade	20,0% (6)	<b>56,7%</b> <b>(17)</b>	10,0% (3)	10,0% (3)	3,3% (1)	30
Experiência com a fonte	10,0% (3)	<b>53,3%</b> <b>(16)</b>	26,7% (8)	10,0% (3)	0,0% (0)	30
Baixo custo financeiro	<b>40,0%</b> <b>(12)</b>	33,3% (10)	16,7% (5)	6,7% (2)	3,3% (1)	30
Língua	6,7% (2)	<b>50,0%</b> <b>(15)</b>	33,3% (10)	6,7% (2)	3,3% (1)	30

Dentre as opções com menos relevância, ao considerar menos importante os valores contidos entre “indiferente” e “sem importância”, a língua (43,3%, ou 13 pessoas) foi a menos considerada, seguida por experiência com a fonte (36,7%, ou 11 pessoas) e baixo custo financeiro (26,7%, ou 8 pessoas).

Em suma, para os inquiridos o aspecto mais relevante no acesso e escolha das fontes de informação electrónicas, foi a obtenção de conteúdo actualizado, fiável, de boa qualidade e relevante para o desenvolvimento do trabalho de investigação.

**Questão 3:** Para o desenvolvimento do trabalho de investigação, indique o grau de incidência das dificuldades – indicados no inquérito – encontradas na pesquisa de informação.

No universo dos 30 inquiridos, a maioria indicou o excesso de informação (66,7%, ou 20 pessoas) como uma dificuldade frequentemente encontrada, seguida por informação indesejada (50%, ou 15 pessoas) e informação indisponível (40%, ou 12 pessoas).

**TABELA 6 – Dificuldades encontradas na pesquisa de informação**

Dificuldades	Sempre	Frequentemente	As vezes	Quase nunca	Nunca	Respostas
Localização da informação	3.3% (1)	33.3% (10)	<b>53.3%</b> <b>(16)</b>	10.0% (3)	0.0% (0)	30
Excesso de informação	6.7% (2)	<b>66.7%</b> <b>(20)</b>	16.7% (5)	10.0% (3)	0.0% (0)	30
Informação indisponível	3.3% (1)	40.0% (12)	<b>50.0%</b> <b>(15)</b>	6.7% (2)	0.0% (0)	30
Informação indesejada	6.7% (2)	<b>50.0%</b> <b>(15)</b>	33.3% (10)	10.0% (3)	0.0% (0)	30
Utilização do motor de pesquisa	0.0% (0)	10.0% (3)	36.7% (11)	<b>43.3%</b> <b>(13)</b>	10.0% (3)	30
Interpretação dos resultados obtidos pelo motor de pesquisa	3.3% (1)	20.0% (6)	<b>33.3%</b> <b>(10)</b>	<b>33.3%</b> <b>(10)</b>	10.0% (3)	30

O excesso de informação normalmente ocorre quando na pesquisa em fontes de informação pouco estruturadas, que, na maioria das vezes, não utilizam nenhuma ferramenta de auxílio para a classificação do conteúdo, como é o exemplo da *Internet* e dos motores de pesquisa.

Em relação aos inquiridos, é compreensível que fosse esta a dificuldade mais evidenciada, pois, como observado na questão 1 desta secção, utilizavam normalmente fontes de informação sem metodologia de organização, como motores de pesquisa da *Web*.

#### **6.2.4 Conclusões**

O presente estudo permitiu avaliar, genericamente, o comportamento dos profissionais ligados ao INESC-Porto, na selecção, organização, classificação e partilha de informações.

Sumariamente, pode-se concluir que:

- Os sistemas formais de organização do conhecimento ainda são – naturalmente – mais conhecidos pelos intervenientes do que a folksonomia. Isto não significa, no entanto, que na prática sejam os mais utilizados. Em termos percentuais, o número de utilizadores da folksonomia é superior ao dos utilizadores dos demais sistemas;
- Organizar informações localmente é ainda o método preferencial dos respondentes, através da estruturação de pastas para guardar documentos e *bookmarking* para *sites* da *Internet*;
- A preferência pela partilha de informações recai sobre os colegas do mesmo departamento/grupo de investigação e, posteriormente, sobre colegas de outras instituições;
- A *Internet* é a principal fonte para a pesquisa de informação no desenvolvimento do trabalho de investigação. Contudo, o processo de pesquisa ocorre ainda de forma não organizada. Para além deste indicador, as publicações externas, os colegas do mesmo grupo de investigação e a biblioteca pessoal são também referenciadas como fontes frequentemente utilizadas;
- A boa qualidade técnica e a actualidade da informação são os factores que mais influenciam na escolha e acesso às fontes, contrastando assim com a preferência pela língua, experiência com a fonte e baixo custo financeiro;
- Apesar das inúmeras vantagens reconhecidas, foram apontadas como principais dificuldades na pesquisa de informação, nomeadamente, excesso de informação, informação indesejada e informação indisponível.

A popularidade dos esquemas de classificação entre a maioria dos intervenientes é um factor naturalmente compreensível. A sua utilização em bibliotecas para a organização do conhecimento é secular. No entanto, e como o conteúdo das bibliotecas é frequentemente utilizado para desenvolvimento de trabalhos técnico-científicos, é natural que, para os seus utilizadores regulares, os esquemas de classificação sejam mais familiares do que outros tipos de classificação.

Em relação ao factor popularidade, o mesmo ocorre com a taxonomia e com a ontologia. É natural que grande parte dos intervenientes conheçam ou já tenham ouvido falar desses SOC, pois, além de estarem maciçamente presentes na literatura científica, também são sistemas bastante utilizados por empresas e instituições. Por serem sistemas flexíveis e adaptáveis às necessidades, por vezes, são melhores opções às organizações do que os tradicionais esquemas de classificação.

Em relação à folksonomia, alguns factores contribuem para que ela seja menos familiar entre os intervenientes do que os esquemas de classificação, taxonomias e ontologias. Mesmo sendo cada vez mais comum em *sites* da *Web* para classificar conteúdos, a folksonomia não é um SOC frequentemente utilizado por empresas e instituições. Além disso, o tema é ainda pouco explorado na literatura científica, apesar de ser um assunto constantemente citado em *blogs*, principalmente naqueles que exploram conteúdos relacionados à *Web2.0*.

No que diz respeito à organização de documentos e conteúdos, apesar da actual diversidade de ferramentas existentes, a maioria dos intervenientes revelam nas suas respostas adoptarem ainda formas mais tradicionais, como pastas do próprio computador. Tais opções diminuem, ou por vezes extinguem, a possibilidade de descentralização de acesso às informações e a sua partilha de forma simultânea, aspectos, que por sua vez, são muito importantes para o desenvolvimento de um trabalho de investigação.

Apesar da avaliação dos factores que influenciam tais escolhas não estarem presentes neste inquérito, podemos considerar algumas prováveis causas, tais como resistência a mudanças, falta de tempo e opção pelo mais usual e familiar, e não pelo mais vantajoso, ou talvez, pelo sentimento de segurança em relação às informações, pois é mais fácil controlar o seu acesso quando elas estão disponíveis apenas no computador pessoal.

Tal comportamento também poderia ser ilustrado através dos resultados obtidos por (Hertzum 2002), que relata que para criar uma relação de familiaridade entre os engenheiros e um sistema de informação, não basta que o sistema esteja acessível em qualquer lugar e a qualquer hora, mas que, para além da acessibilidade, é necessário que ele seja familiar e que a sua veracidade seja fácil de determinar. No entanto, é necessário uma análise mais profun-



da desta matéria, com base num estudo mais detalhado, para perceber as reais motivações dos utilizadores das diferentes ferramentas.

Semelhante aos resultados obtidos por (Ellis and Haugan 1997), (Anderson, Glassman et al. 2001), (Hertzum and Pejtersen 2000), (Kwasitsu 2003), (Fidel and Green 2004), os colegas do mesmo departamento ou grupo de investigação são, para além de fontes de informação frequentemente utilizadas, também pessoas com quem mais partilham informações. Uma razão a ser considerada como motivação para tal facto, poderia ser a proximidade física. Mas, segundo (Fidel and Green 2004), este não é o principal factor de motivação, e sim a familiaridade, pois é ela que contribui para minimizar o esforço intelectual na selecção de informações fiáveis. No entanto, aqui também é necessário o desenvolvimento de um estudo mais detalhado para perceber quais são os principais factores que motivam tais escolhas.

Da mesma maneira que foi observado no trabalho de (Hirsh and Dinkelacker 2004), o comportamento de pesquisa da maioria dos intervenientes foi dominado pelo uso da *Internet*, principalmente de uma forma não organizada. Os colegas de trabalho do mesmo grupo de investigação e a biblioteca pessoal também foram consideradas fontes de informação frequentemente utilizadas. Isto aponta para o facto de que o perfil de grande parte dos inquiridos incidiu pela escolha das fontes de informação mais acessíveis e, em muitos casos, mais rápidas na resposta a um problema, independentemente da precisão dos resultados apresentados. Citando como exemplo, uma pesquisa realizada numa fonte de informação estruturada, como Metalib, B-On, Compedex, Inspec etc., é muito mais precisa do que a mesma pesquisa realizada numa fonte de informação não estruturada, como o Google. Independentemente deste factor, constatou-se ser muito maior a preferência por motores de pesquisa do que por outras fontes de informação.

A preferência por pesquisas em fontes de informação não estruturadas, talvez seja a principal causa para o excesso de informação, informação indesejada e informação indisponível, factores estes, referidos pelos intervenientes como dificuldades mais encontradas no processo de pesquisa.

Apesar do número de respostas obtidas por este questionário ter ficado aquém das expectativas, ainda assim, permitiu delinear o perfil dos profissionais, objecto deste estudo. Sucintamente, os mesmos tendem a optar por processos que lhe são mais familiares – como por exemplo, organização de conteúdos em pastas do computador pessoal e obtenção de informações junto a colegas do mesmo grupo de investigação –, preterindo processos, por vezes, mais eficazes, como a partilha de informação através da intranet ou *wikis*.

Deste estudo ressaltam pois conclusões importantes que definem o comportamento na pesquisa de informação. Apesar das inúmeras fontes existentes capazes de suprir as necessida-

des dos seus utilizadores, a forma não estruturada da pesquisa de informação contribui para as maiores dificuldades encontradas no processo, como por exemplo, o excesso de informação.

## **7 PROPOSTA DE UMA COMBINAÇÃO DE ESTRATÉGIAS *TOP-DOWN* E *BOTTOM-UP* NA GESTÃO DE INFORMAÇÃO DO INESC-PORTO**

Como base nos capítulos anteriores, pretende-se aqui explorar uma possível forma de utilização de estratégias de classificação estruturadas e não estruturadas para a gestão de informação do INESC-Porto.

Para além da preocupação com a eficiência e eficácia decorrentes da aplicação desses sistemas, o factor “uso” foi de fundamental importância para se considerar qualquer opção. Isto quer dizer que, mesmo que teoricamente um sistema fosse melhor do que outro, caso ele tivesse grande probabilidade de não ter continuidade pelos utilizadores, seria desconsiderado.

Entretanto, os tradicionais sistemas de classificação, como, por exemplo, os esquemas universais, não foram considerados nesta abordagem, pois, apesar de eficientes, neste caso, não eram eficazes, devido à sua dificuldade em incorporar rapidamente novos assuntos, uma constante na área de I&D. Com o passar do tempo, tal facto certamente os tornaria sistemas inapropriados e, conseqüentemente, seriam descontinuados pelos utilizadores. No caso dos tesouros, vocabulário controlado e cabeçalho de assuntos, foram considerados, apenas, como possíveis ferramentas para auxílio na classificação.

No caso do INESC-Porto, pelo menos numa fase inicial, não será necessário associar definições e conceitos aos conteúdos e objectos contidos no sistema de classificação. Com isto, não foi também considerada a ontologia, pois, com a excepção para estas funcionalidades de associação, a taxonomia desenvolve um papel semelhante ao seu.

Como observado, os esquemas de classificação aqui ponderados foram a taxonomia e a folksonomia. Tal selecção foi baseada, principalmente, nas características correspondentes à flexibilidade e capacidade na rápida absorção de novos temas. Na área de I&D tais factores são fundamentais para os sistemas de gestão de informação, pois esta é uma área em constante transformação.

A secção 6.2, onde se avaliou o comportamento de selecção, organização, classificação e partilha de informação dos profissionais ligados ao INESC-Porto, serviu de base para todas as considerações decorrentes neste capítulo.

Como demonstrado anteriormente, os profissionais que actuam nesse instituto escolhem e acedem a informações baseadas, sobretudo, na sua actualidade e boa qualidade técnica. No entanto, para eles, a principal forma de recuperar informação é através de pesquisas não

estruturadas na *Internet*. Isto é certamente, uma das causas para as suas principais dificuldades na pesquisa por informação, nomeadamente, excesso de informação, informação indesejada e informação indisponível.

Outro factor observado foi que esses profissionais, além de considerarem colegas do mesmo grupo de investigação como fontes de informação frequentemente utilizadas, também indicaram que, para além deles, colegas externos ao instituto eram indivíduos com quem normalmente partilhavam informações. Isto prova que não só os materiais impressos eram considerados recursos bons e fiáveis para a recuperação de informação, mas também as pessoas da sua confiança.

Em relação à organização de documentos e conteúdos, eles ainda preferiam fazê-lo localmente, tal como em pastas do próprio computador. Da análise realizada no capítulo anterior, algumas possíveis causas para este comportamento prendiam-se sobretudo com os factores resistência à mudança, falta de tempo, opção pelo mais usual e familiar em vez do mais vantajoso, e segurança das informações. Apesar destas motivações não terem sido estudadas neste trabalho, tais factores foram considerados no presente capítulo.

Do conjunto das informações acima explicitadas, é possível mencionar alguns factores a ter em consideração para a implementação de um sistema de gestão de informação no INESC-Porto, que se revele útil e atractivo aos utilizadores, tais como:

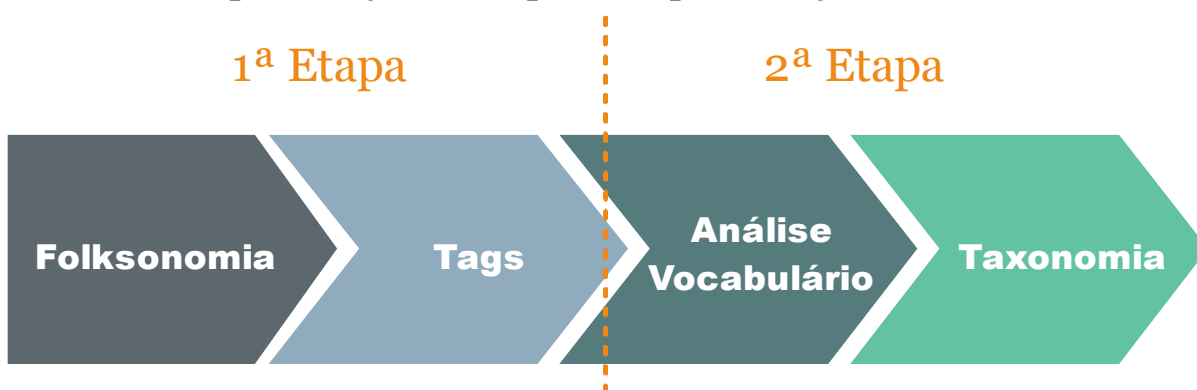
- Fácil e intuitivo, por forma a permitir a maximização do tempo na utilização e o impacto na rotina dos profissionais;
- A garantia que nos resultados apresentados apareçam informações actuais e de boa qualidade, ou que, pelo menos, seja indicada a procedência da informação para que o próprio utilizador faça este julgamento;
- A similaridade com os demais sistemas de pesquisa existentes na Web, ou então, a capacidade de se transformar rapidamente em algo usual e familiar;
- A diminuição do excesso de informação e informações irrelevantes devolvidas no resultado de uma pesquisa;
- A possibilidade de considerar e permitir a classificação de indivíduos como recursos de informação;
- A manutenção da confidencialidade e segurança das informações, sempre que se revele necessário.

De acordo com estas características, nenhum dos sistemas aqui considerados – taxonomia e folksonomia – abrange por completo tais necessidades. Por exemplo, uma taxonomia poderia garantir resultados com informações mais bem direccionadas e mais relevantes, o que diminuiria consideravelmente o excesso de informação, mas, por mais simples que fosse, exigiria que os seus utilizadores tivessem algum esforço cognitivo para a manusear. Tal factor, poderia, por vezes, diminuir o interesse pelo uso do sistema, tanto pelo dispêndio de tempo quanto pela falta de usabilidade.

Por outro lado, a folksonomia não garante resultados sem excesso de informação. Contudo, ela é um sistema simples e intuitivo, que não exige quase nenhum esforço cognitivo dos utilizadores, mas apenas que eles atribuam termos próprios na classificação dos conteúdos para a sua recuperação futura. No entanto, para além disso, este sistema permite que seja feita uma partilha de informação sem grande esforço para os seus utilizadores.

Ao contrário do exemplo apresentado por (Gibson 2005), onde existia uma taxonomia e em seguida foi implementada uma folksonomia, este estudo pretende propor uma combinação de estratégias de classificação implementadas em duas etapas, sendo: a) um sistema de folksonomia; b) a partir das informações obtidas através da folksonomia, desenvolver e implementar um sistema taxonómico. Tal procedimento é semelhante ao caso do STARLab (Christiaens 2006), onde uma folksonomia originou uma taxonomia e, conseqüentemente, uma ontologia.

**FIGURA 7 – Representação das etapas de implementação**



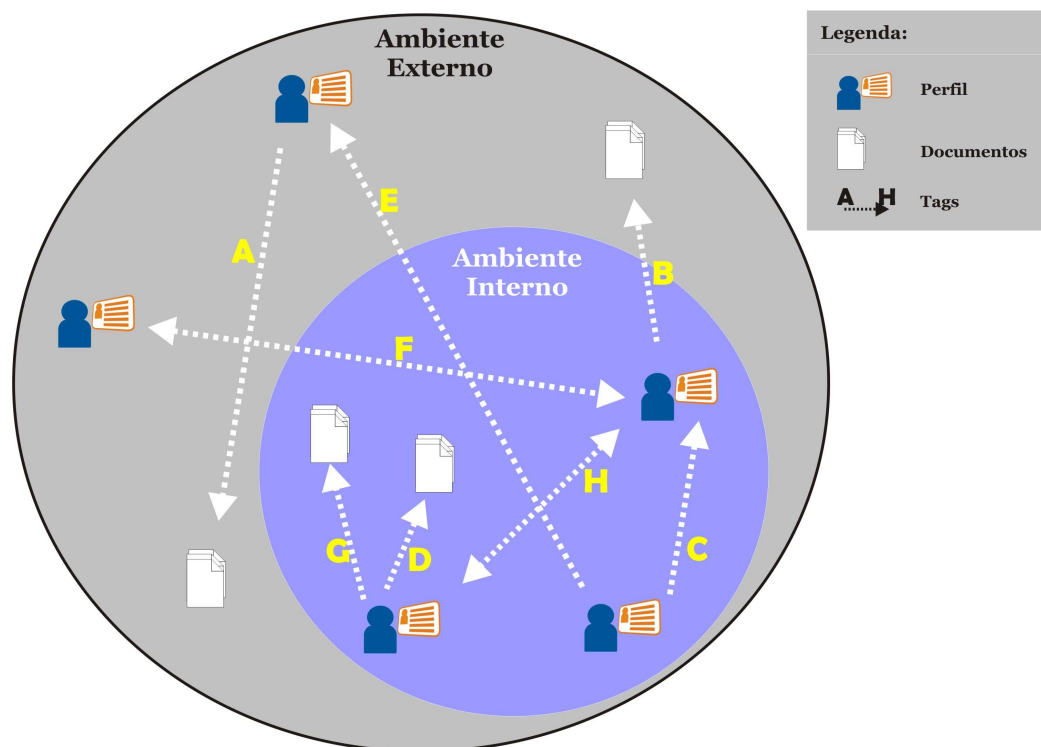
A proposta de implementação em etapas diferentes tem por objectivo criar um processo gradativo e que interfira o menos possível na rotina dos profissionais. Desta forma, a probabilidade deste procedimento ter mais receptividade e de se tornar realmente uma ferramenta que acrescente valor ao instituto, é consideravelmente maior.

A escolha pela implementação da folksonomia numa primeira fase dá-se pelo facto de ser um sistema de simples implementação e utilização. Além disso, ela oferece uma maneira de observar o comportamento de classificação dos seus utilizadores, identificando os termos mais usados para um tipo de documento, conteúdo, recurso etc. Isto fornece elementos para que seja criada, em seguida, uma estrutura bem organizada à semelhança das necessidades dos profissionais do instituto, ou seja, uma taxonomia.

O ideal é que a folksonomia tenha algumas funcionalidades semelhantes ao Dogear, da IBM, mencionado na secção 4.2.2. A primeira será que os utilizadores, ao fazerem o registo no sistema, adoptem a utilização de nomes reais para facilitar a comunicação entre eles, como também para auxiliar na avaliação de informações através dos indivíduos que estão associados a ela.

Para além desta funcionalidade, uma outra, não menos importante, será permitir que, para além das informações, sejam incluídas tags ao perfil dos utilizadores, como ocorre no BluePages+1 (secção 4.2.3). Isto possibilitará que cada indivíduo classifique o perfil dos profissionais da forma mais conveniente ao seu trabalho.

**FIGURA 8 – Representação do esquema folksonómico**



Ainda similarmente ao Dogear, outra funcionalidade será incluir sistemas de alerta que avise os utilizadores quando novas informações forem incluídas. Tal funcionalidade estimulará o uso do sistema, e fará com que o processo de pesquisa e partilha de informação se torne mais rápido e proveitoso.

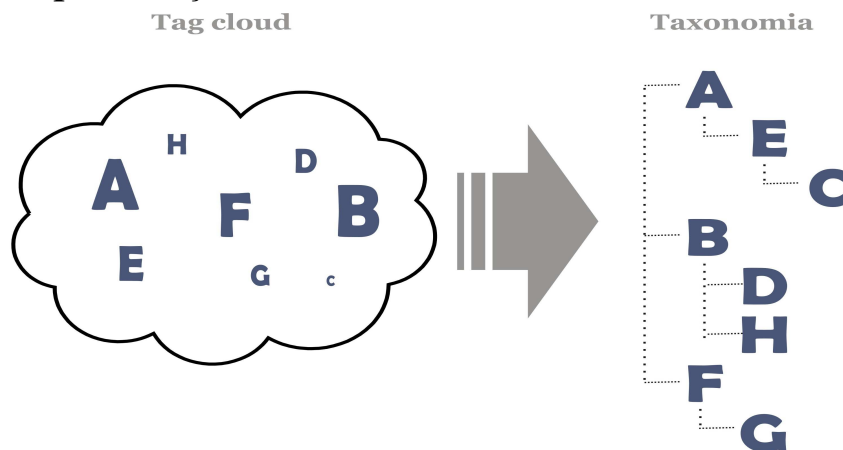
A partir destes processos, seria possível obter, desde logo, alguns benefícios, tais como:

- Classificar informações de forma personalizada, o que resultaria numa rápida recuperação;
- “Emprestar” termos atribuídos por outros utilizadores;
- Pesquisar de forma rápida e simples, como ocorre na Web. No entanto, com uma vantagem acrescida, a possibilidade de encontrar conteúdos sem estar à espera, bastando apenas navegar através dos termos atribuídos por outros utilizadores;
- Encontrar e comunicar-se rapidamente com profissionais da rede de trabalho;
- Avaliar uma informação através do seu conteúdo ou através dos utilizadores a ela associados;
- Estar sempre actualizado em relação ao conteúdo existente no sistema;
- Classificar e partilhar informação (muitas vezes involuntariamente) praticamente em tempo real;
- Controlar o conteúdo existente (de acordo com as permissões de acesso).

Apesar da folksonomia prover tais benefícios, fica um pouco aquém quando se trata de uma gestão de informação estruturada. Ao contrário da taxonomia, ela não diminui o excesso de informação obtido numa pesquisa e nem organiza os temas numa estrutura hierárquica. Do ponto de vista geral, um sistema folksonómico, por vezes, pode até aumentar o excesso de informação, mas, do ponto de vista individual, ele pode ser uma boa opção para pesquisas rápidas.

No entanto, para suprir essas lacunas deixadas pela folksonomia, relacionadas com a organização estruturada de informação, propõe-se, num segundo momento, a implementação de uma taxonomia à luz dos resultados obtidos, através da utilização e avaliação da classificação feita a partir do sistema folksonómico.

**FIGURA 9 – Representação da taxonomia baseada na folksonomia**



Um factor a ser considerado no processo de implementação tanto da folksonomia quanto da taxonomia, é a segurança e a confidencialidade das informações. O sistema deverá ser seguro de forma a permitir o acesso às informações somente aos utilizadores autorizados, sejam eles internos ou externos ao instituto.

É importante ressaltar que ambos os sistemas deverão estar totalmente integrados para que a gestão de informação do INESC-Porto seja feita de uma maneira otimizada, de forma a tornar-se uma mais-valia no processo de I&D.

Através desta integração, os utilizadores terão três opções para fazer as suas pesquisas dos conteúdos disponíveis no instituto: a) no caso de pesquisas mais complexas e precisas, utilizarão a taxonomia para pesquisar de uma forma estruturada; b) através da folksonomia, no caso de pesquisas mais rápidas e simples; c) ou através de uma combinação de ambas as formas, podendo interagir entre os dois sistemas, movendo-se da taxonomia para as *tags* e vice-versa.

Por fim, além dos benefícios mencionados anteriormente, a existência de um sistema de gestão de informação na intranet do INESC-Porto, facilitará o acesso às informações a partir de qualquer computador ligado à intranet ou Internet. No trabalho de I&D, a eliminação das barreiras tempo e espaço é um factor de extrema importância, pois permite que os profissionais acedam aos seus documentos e informações independentemente da hora e do local.

Em síntese, através da implementação de tais sistemas, acredita-se que a gestão de informação do INESC-Porto tornar-se-á cada vez mais eficiente e eficaz, pois o processo de organização que até então era realizado de maneira individual, será efectuado de forma integrada e colaborativa, bem como as limitações de espaço e tempo para acesso às informações virão a ser reduzidas, ou, até eliminadas. Outra vantagem é que todo este processo não acarretará



grandes alterações na rotina dos profissionais, pois, a prática exercida por eles nos próprios computadores passará a ser efectuada de maneira semelhante num sistema disponível na rede interna do instituto.

Com base na exposição feita neste capítulo, um dos sistemas que poderia ser adoptado é o Lotus Connections da IBM (IBM 2008). Esta plataforma disponibiliza cinco principais serviços, sendo eles:

- **Perfis:** guarda informações gerais sobre contactos, com a possibilidade de atribuição de *tags* aos perfis de interesse;
- **Comunidades:** possibilita a reunião de utilizadores acerca de uma área de interesse, disponibiliza fóruns online e grupos de discussão, *bookmarks* partilhados e *new feeds*. Além disso, oferece funcionalidades avançadas para wiki;
- **Blogs:** permite que cada indivíduo que tenha um perfil no sistema crie um blog;
- **Dogear:** serviço de *social bookmarking* que permite aos utilizadores criar bookmarks para conteúdos internos ou externos à instituição e partilhá-los de forma segura com outros utilizadores do sistema;
- **Actividades:** serviço que fornece um único destino Web para armazenar lista de tarefas a realizar, ficheiros, bookmarks e mensagens para uma equipa. Além disso, disponibiliza tipos de *feeds* para alterações e actualizações.

As características funcionais deste sistema enquadram-se perfeitamente nas definições anteriormente expostas no decorrer deste capítulo, daí a opção pela sua indicação em particular. Nesta escolha não foram tidos em consideração os requisitos mínimos do sistema e nem o seu custo de implementação e manutenção.

No entanto, a implementação deste sistema numa instituição como o INESC-Porto requer, obviamente, um estudo prévio sobre todas as variáveis que possam influenciar na sua adopção, sejam elas funcionais, técnicas ou financeiras.

## 8 CONCLUSÕES E TRABALHOS FUTUROS

### 8.1 Conclusões

O presente estudo teve como objectivo melhorar a eficiência e eficácia na gestão de informação em instituições de I&D, através da utilização de uma combinação de estratégias de classificação *top-down* e *bottom-up*.

Mais especificamente, com base num estudo – aqui realizado – sobre o comportamento de selecção, organização, classificação e partilha de informação dos profissionais que actuam no INESC-Porto, pretendeu-se delinear uma estratégia de classificação de informação segundo uma combinação de SOC formais e informais.

Para a obtenção destes resultados, a metodologia adoptada verteu sobre os seguintes parâmetros:

- Revisão de literatura e compilação das características, vantagens e desvantagens dos SOC (formais e informais);
- Compilação de trabalhos relacionados com a comparação entre SOC;
- Realização de um estudo comparativo, com base nos dois tópicos anteriores, entre a folksonomia e alguns desses sistemas;
- Revisão de literatura sobre o comportamento de pesquisa de informação dos engenheiros e cientistas;
- Realização de um estudo de caso realizado no INESC-Porto sobre o comportamento de selecção, organização, classificação e partilha de informação dos profissionais que actuam, directa ou indirectamente, no instituto;
- Apresentação de uma proposta de estratégia de classificação para a gestão de informação do instituto, a partir da combinação das informações obtidas nos tópicos anteriores.

Durante todo o processo de investigação foram identificadas duas principais dificuldades. A primeira relacionada com a literatura científica existente em relação à folksonomia, mais especificamente sobre a folksonomia empresarial. A segunda relacionada com a recolha das respostas do questionário aplicado aos profissionais do INESC-Porto.

A folksonomia é um tema sobejamente abordado em sites da *Web*, e cada vez mais explorado – nos mais diversos contextos – na literatura científica. No entanto, a folksonomia empresarial não é ainda um tema muito explorado na área académica. Tal factor poderá ser explicado porque eventualmente são ainda poucas as empresas que a utilizam para gerir as próprias informações. Um exemplo de uma instituição que tem apostado nesta forma de organização – juntamente com a combinação de sistemas formais – é a IBM. Para além da prática, ela investe igualmente na investigação e divulgação de materiais relacionados com o tema.

Versando sobre o estudo de caso, foram registadas algumas dificuldades na obtenção das respostas ao questionário remetido. Como anteriormente explicitado no capítulo 6, apenas uma parte dos visados responderam ao inquérito. É possível, desde logo, identificar uma principal razão, mais concretamente a data em que o inquérito foi lançado, Julho e Agosto, período que colidiu, na maioria dos casos, com a ausência para férias dos profissionais envolvidos. Contudo, apesar das circunstâncias, apenas neste momento era oportuno lançar este inquérito, tendo em conta a evolução do trabalho de investigação e no respeito absoluto pelos prazos definidos para a conclusão deste projecto.

Tendo em consideração a implementação dos SOC, pode constatar-se que o ambiente no qual ele será aplicado é um factor decisivo para a escolha da tipologia do sistema ou, por outro lado, talvez possamos referir que a melhor escolha é uma combinação entre dois ou mais sistemas de classificação.

Tendo esta noção como base, o estudo de caso realizado no INESC-Porto identificou algumas características que permitiram delinear, para a sua gestão de informação, a possibilidade de implementar um sistema através da combinação de estratégias de classificação formais e informais.

Em relação ao ambiente, verificou-se que os profissionais que actuam no instituto têm por hábito pesquisar informação, maioritariamente, nos canais informais e não estruturados e a partilha é realizada entre colegas de trabalho e colegas externos ao instituto. No entanto, quando se trata da organização e classificação de informações, existe ainda um certo conservadorismo, pois a opção recai sobre segurança e facilidade, ou, comodidade, dando preferência ao computador pessoal, preterindo um sistema partilhado.

A partir da constatação desses comportamentos, pôde aferir-se que uma boa opção para a gestão de informação do INESC-Porto seria a combinação entre uma folksonomia e uma taxonomia. Para além disto, sugeriu-se uma implementação em duas etapas diferentes, permitindo, assim, que o sistema fosse bem aceite e utilizado pelos profissionais. Num primeiro momento, optou-se pela implementação de uma folksonomia, pelo facto de ser um sistema simples e intuitivo, com pouca interferência na rotina dos profissionais envolvidos na investi-

gação, mas passível de criar importantes resultados. Numa segunda fase, para permitir uma organização estruturada do conteúdo disponível no sistema em questão, foi proposta a implementação de uma taxonomia combinada com a folksonomia.

É com base neste estudo elaborado, tema deste projecto de investigação, que se pode formular a opinião de que em ambientes de I&D é extremamente importante, para a optimização da gestão de informação, o recurso a um sistema que possa incorporar canais formais (ex.: livros, artigos, documentos etc.) e informais (ex.: colegas de trabalho). Caberá depois aos utilizadores do sistema implementado a sua melhor gestão para que se obtenham os melhores proveitos face às suas necessidades específicas.

Apesar dos resultados apurados neste trabalho de investigação e consequentes conclusões retiradas, este tema é ainda susceptível de novas abordagens, a partir de diferentes perspectivas, para que seja possível obter uma visão abrangente das necessidades e motivações comportamentais dos profissionais que actuam no INESC-Porto, e paralelamente estudar a metodologia de implementação de um sistema que melhor se adapte à instituição.

## ***8.2 Trabalhos Futuros***

Os resultados obtidos neste estudo apresentaram apenas um pequeno passo em direcção ao melhor aproveitamento e optimização da gestão de informação do INESC-Porto. É necessário, contudo, a realização de novos trabalhos para se obter uma perspectiva mais detalhada, tanto do ponto de vista comportamental, quanto do ponto de vista técnico.

Uma perspectiva não abordada neste estudo, mas que também é de fundamental importância, está relacionada às motivações que levam os profissionais do INESC-Porto a terem determinados comportamentos ligados à selecção, organização, classificação e partilha de informação. O facto é que não basta somente saber o que fazem estes profissionais, mas também saber o que pensam, pois só assim é possível delinear um sistema que realmente satisfaça as suas necessidades em relação à gestão de informação.

Além disso, as características técnicas relacionadas tanto com a escolha, quanto com a implementação de cada sistema devem ser cautelosamente estudadas. Importa definir quais serão os tipos de sistemas a serem implementados, mas também quais os procedimentos a considerar. Torna-se fundamental aferir se a melhor opção é por sistemas já existentes (gratuitos mas com menos funcionalidades, ou pagos e completos), ou então, um sistema desenvolvido à medida. No entanto, em qualquer um dos casos, deve-se ter em conta a melhor res-

posta às necessidades dos profissionais do instituto e a possibilidade de integração com outros sistemas.

No caso da segurança das informações, é importante que o sistema tenha um controlo de acesso fiável, mas que, ao mesmo tempo, não se torne demasiadamente “burocrático” e desinteressante para os seus utilizadores. No entanto, a definição de uma boa política de segurança depende de um estudo que permita definir qual a melhor opção a ser implementada.

Como proposto no capítulo 7, uma combinação entre a folksonomia e a taxonomia deve ser feita a partir dos resultados obtidos através da folksonomia. Assim, antes da implementação de um sistema taxonómico, que deverá ter completa interacção com o outro sistema, deve ser feita uma análise entre os objectos marcados pela folksonomia e as *tags* utilizadas neste processo. A partir desta análise um estudo deve ser realizado de forma a organizar hierarquicamente os temas identificados no processo anterior, bem como uma análise que possa identificar a melhor forma para integrar a taxonomia no sistema já existente.

Por fim, caso se verifique necessário ou interessante ao processo de I&D do INESC-Porto, podem avaliar-se possíveis formas de, a partir do sistema folksonómico e taxonómico, agregar mais um sistema que descreva e defina os conceitos dos objectos existentes no sistema de gestão de informação do instituto, ou seja, uma ontologia.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Al-Khalifa, H. S. (2007). Automatic document-level semantic metadata annotation using folksonomies and domain ontologies. United Kingdom, Faculty of Engineering and Applied Science, School of Electronics and Computer Science, University of Southampton.
- Alcatel-Lucent. (2008). "Alcatel-Lucent." Retrieved March 11th, 2008, from <http://www.alcatel-lucent.com>.
- Anderson, C. J., M. Glassman, et al. (2001). "An investigation of factors affecting how engineers and scientists seek information." Journal of Engineering and Technology Management **18**(2): 131-155.
- Angeles, M. (2005). Making libraries more delicious: social bookmarking in the enterprise.
- Auray, N. (2007). Folksonomy: the new way to Serendipity.
- Beck, H. and H. S. Pinto (2002). Overview of approach, methodologies, standards, and tools for ontologies, University of Florida: The Agricultural Ontology Service (UN FAO).
- Bielenberg, K. and M. Zacher (2005). Groups in social software: utilizing tagging to integrate individual contexts for social navigation, Universität Bremen. Program of Digital Media. Master of Science in Digital Media.
- Braly, M. D. and G. B. Froh (2006). Social Bookmarking in the Enterprise.
- Britannica, E. (2008). taxonomy. Encyclopædia Britannica Online. E. Britannica, Encyclopædia Britannica.
- Broughton, V., J. Hansson, et al. (2005). European curriculum reflections on library and information science education. L. Kajberg and L. Lørring, Royal School of Library and Information Science, The: 133-148.
- Brouwers, M., A. Nawaz, et al. (2007). Folksonomies.
- Butterfield, S. (2004). I am sharing this with you.
- Caplan, P. (2003). Metadata Fundamentals for All Librarians, ALA Editions.
- Choo, C. W. (1996). "The knowing organization: How organizations use information to construct meaning, create knowledge and make decisions." International Journal of Information Management **16**(5): 329-340.
- Christiaens, S. (2006). Metadata mechanisms: from ontology to folksonomy ... and back, Semantics Technology and Applications Research Laboratory, Vrije Universiteit Brussel.
- Cook, N. (2006). The case for enterprise bookmarking.

- Cook, N. (2007). Fifteen Uses of Corporate Bookmarking.
- Corazzon, R. (2008). Ontology. A Resource Guide for Philosophers.
- Cormier, P. (2006). Social tagging for the enterprise.
- CREMONI, L. (2008). "Web 2.0 map." RailsOnWave Retrieved February 4th, 2008, from <http://www.railsonwave.it/railsonwave/2007/1/2/web-2-0-map>.
- Damme, C. V. (2006). Folksonomies and Enterprise Folksonomies, Faculty of Economic, Political and Social Science and Solvay Management School.
- Davenport, T. H. and L. Prusak (2000). "Working knowledge: how organizations manage what they know." Ubiquity **1**(24): 2.
- Delphi, G. (2002). Taxonomy & Content Classification.
- Dye, J. (2006). "Folksonomy: A Game of High-tech (and High-stakes) Tag." EContent **29**(3): 38-43.
- Eisner, S. (2006). Our experience with enterprise tagging terminology.
- Ellis, D. and M. Haugan (1997). "Modelling the information seeking patterns of engineers and research scientists in an industrial environment." Journal of Documentation **53**(4): 384-403.
- Eppler, M. J. and J. Mengis (2003). A framework for information overload research in organizations: insights from organization science, accounting, marketing, MIS, and related disciplines.
- Fettke, P. and P. Loos (2003). "Classification of reference models: a methodology and its application." Information Systems and E-Business Management **1**(1): 35-53.
- Fidel, R. and M. Green (2004). "The many faces of accessibility: engineers' perception of information sources." Information Processing & Management **40**(3): 563-581.
- Fitzgerald, M. (2007). Tagging tools offer powerful way to organize information.
- Garshol, L. M. (2004). "Metadata? Thesauri? Taxonomies? Topic maps! Making sense of it all." Journal of Information Science **30**(4): 378--391.
- Gibson, B. (2005). IBM's Intranet and Folksonomy.
- Gilchrist, A. (2001). "Corporate taxonomies: report on a survey of current practice." Online Information Review **25**(2): 94-102.
- Gilchrist, A. (2003). "Thesauri, taxonomies and ontologies – an etymological note." Journal of Documentation **59**(1): 7-18.

- Goffee, R. and G. Jones (1996). "What Holds the Modern Company Together?" Harvard Business Review **74**(6): 133-148.
- Golder, S. A. and B. A. Huberman (2006). "Usage patterns of collaborative tagging systems." Journal of Information Science **32**(2): 198-208.
- Gruber, T. R. (1993). "A translation approach to portable ontology specifications." Knowledge Acquisition **5**(2): 199-220.
- Guo, B. and H. Li (2007). Information seeking behavior of R&D professionals in new product development. Industrial Engineering and Engineering Management, IEEE.
- Guy, M. and E. Tonkin (2006). "Folksonomies: Tidying Up Tags?" D-Lib Magazine **12**(1): web.
- Hammond, T., T. Hannay, et al. (2005). "Social bookmarking tools (I): a general review." D-Lib Magazine **11**(4): web.
- Harris, T. G. (1993). "The Post-Capitalist Executive: An Interview with Peter F. Drucker." Harvard Business Review **71**(3): p114 - 122.
- Hayman, S. (2007). Folksonomies and tagging: new developments in social bookmarking.
- Hertzum, M. (2002). "The importance of trust in software engineers' assessment and choice of information sources." Information and Organization **12**(1): 1-18.
- Hertzum, M. (2008). "Collaborative information seeking: the combined activity of information seeking and collaborative grounding." Information Processing & Management **44**(2): 957-962.
- Hertzum, M. and A. M. Pejtersen (2000). "The information-seeking practices of engineers: searching for documents as well as for people." Information Processing & Management **36**(1): 761-778.
- Hirsh, S. and J. Dinkelacker (2004). "Seeking information in order to produce information: An empirical study at Hewlett Packard Labs." Journal of the American Society for Information Science and Technology **55**(9): 807-817.
- Hodge, G. (2000). Systems of knowledge organization for digital libraries: beyond traditional authority files, The Digital Library Federation, Council on Library and Information Resources.
- IBM. (2008). "IBM Lotus Connections." from <http://www-01.ibm.com/software/lotus/products/connections/>.
- INESC-Porto. (2008). "Apresentação do INESC Porto." Retrieved 10-12-2007, 2008, from <http://www2.inescporto.pt/apresentacao>.



- Infopédia. (2003-2008, September 14th, 2008). "análise SWOT." Infopédia, from [http://www.infopedia.pt/\\$analise-swot](http://www.infopedia.pt/$analise-swot).
- Koch, T., A. Brümmer, et al. (1997). The role of classification schemes in Internet resource description and discovery, DESIRE - Development of a European Service for Information on Research and Education.
- Kremer, S., L. M. Kolbe, et al. (2005). "Towards a procedure model in terminology management." Journal of Documentation **61**(2): 281-295.
- Kroski, E. (2005). The hive mind: folksonomies and user-based tagging.
- Kwasitsu, L. (2003). "Information-seeking behavior of design, process, and manufacturing engineers." Library & Information Science Research **25**(4): 459-476.
- Kwasnik, B. H. (1999). "The role of classification in knowledge representation and discovery." Library Trends **48**(1): p22-47.
- Losee, R. M. (2006). "Browsing mixed structured and unstructured data." Information Processing & Management **42**(2): 440--452.
- Mathes, A. (2004). Folksonomies - Cooperative Classification and Communication Through Shared Metadata, Computer Mediated Communication - LIS590CMC, School of Library and Information Science, University of Illinois Urbana-Champaign.
- Merholz, P. (2004). Ethnoclassification and vernacular vocabularies.
- Millen, D. R. and J. Feinberg (2006). Using social tagging to improve social navigation. AHO6 Workshop: Social Navigation and Community-Based Adaptation Technologies, Dublin, Ireland.
- Millen, D. R., J. Feinberg, et al. (2006). Dogear: Social bookmarking in the enterprise. Conference on Human Factors in Computing Systems: Proceedings of the SIGCHI conference on Human Factors in computing systems, Montréal, Québec, Canada, ACM Press.
- Muller, M., W. Geyer, et al. (2007). Exploratory Tag-Based Search in Multiple Enterprise Domains with the Malibu Productivity Assistant. ESI 2007 Workshop.
- Muller, M. J. (2007). Comparing tagging vocabularies among four enterprise tag-based services. GROUP '07: Proceedings of the 2007 international ACM conference on Supporting group work, New York, NY, USA, ACM.
- Noruzi, A. (2007). "Folksonomies: Why do we need controlled vocabulary?" Webology **4**(2): Editorial.
- Pomerantz, J. (2003). Question taxonomies for digital reference, Syracuse University, School of Information Studies.

- Quintarelli, E. (2005). Folksonomies: power to the people. ISKO Italy-UniMIB meeting, Milan.
- Realcom, A. (2008). "AskMe Realcom." Retrieved March 12th, 2008, from <http://www.askmecorp.com/company/default.asp>.
- Reitz, J. M. (2007). controlled vocabulary. Online Dictionary for Library and Information Science, Libraries Unlimited.
- Rosenfeld, L. (2005). Folksonomies? How about metadata ecologies?
- Shirky, C. (2005). Folksonomies + controlled vocabularies.
- Shirky, C. (2005). Ontology is Overrated: Categories, Links, and Tags.
- Sowa, J. F. (2005). Building, Sharing, and Merging Ontologies.
- Star, S. L. and J. R. Griesemer (1989). "Institutional ecology, 'translations' and boundary objects: amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39." Social Studies of Science **19**(3): 387-420.
- Su, C. (2005). An exploration of redundancy in classification schemes on the Web: a study of EIA and Amazon websites, University of North Carolina, School of Information and Library Science.
- Transportation, U. S. D. o. (2008). "TRIS Online." Retrieved February 2th 2008, 2008, from <http://ntlsearch.bts.gov/tris/index.do>.
- Tristão, A. M. D., G. R. B. Fachin, et al. (2004). "Sistemas de classificação facetados e tesouros: instrumentos para organização do conhecimento." Ciência da Informação **33**(2): 161-171.
- Tudhope, D. and M. L. Nielsen (2006). "Introduction to knowledge organization systems and services." New Review of Hypermedia and Multimedia **12**(1): 3-9.
- Uddin, M. N., M. Mezbah-ul-Islam, et al. (2006). "Information description and discovery method using classification structures in Web." Malaysian Journal of Library and Information Science **11**(2): 1-20.
- van Rees, R. (2003). Clarity in the usage of the terms ontology, taxonomy and classification. The 20th CIB W78 Conference on Information Technology in Construction, Waiheke Island, Auckland, New Zealand.
- Vanderlei, T. A. (2006). Um mecanismo de classificação cooperativa para busca e recuperação de componentes de software. Recife, Centro de Informática, Universidade Federal de Pernambuco.
- Vickery, B. (2008). "On 'knowledge organization'." Retrieved November 1th, 2008, from <http://www.lucis.me.uk/knowlorg.htm#start>.

- Wal, T. V. (2005). Explaining and Showing Broad and Narrow Folksonomies.
- Wal, T. V. (2006). Online Information Folksonomy Presentation Posted.
- Wal, T. V. (2007). Folksonomy Coinage and Definition.
- Wikipedia. (2007). "Folksonomy." Retrieved October 4th, 2007, from <http://en.wikipedia.org/wiki/Folksonomy>.
- Wikipedia. (2007). "Tag cloud." Retrieved October 10th, 2007, from [http://en.wikipedia.org/wiki/Tag\\_cloud](http://en.wikipedia.org/wiki/Tag_cloud).
- Wikipedia. (2008). "Publishing." from <http://en.wikipedia.org/wiki/Publishing>.
- Wikipedia. (2008). "Serendipidade." Retrieved February 10th, 2008, from <http://pt.wikipedia.org/wiki/Serendipismo>.
- Wikipedia. (2008). "Social bookmarking." from [http://en.wikipedia.org/wiki/Social\\_bookmarking](http://en.wikipedia.org/wiki/Social_bookmarking).
- Wiktionary. (2008). "Classification." Retrieved 19th March, 2008, from <http://en.wiktionary.org/wiki/classification>.
- Wilson, T. D. (2000). "Human Information Behavior." Information Science Research **3**(2): 49-55.
- Zurek, B. (2006). It's a tag, tag, tag World! Enterprise Next.

## **ANEXO A**

### Inquérito online

Inquérito relativo ao comportamento de selecção, organização, classificação e partilha de informação, disponibilizado numa página *Web* e o seu endereço electrónico enviado por e-mail aos responsáveis de projecto, investigadores, bolseiros e todos os demais colaboradores do INESC-Porto no exercício da sua actividade técnico-científica.

# COMPORTAMENTOS NA SELECÇÃO, ORGANIZAÇÃO E CLASSIFICAÇÃO E PARTILHA DE INFORMAÇÃO

## 1 APRESENTAÇÃO

Apresentação e objectivos do questionário:

Este questionário tem como objectivo a análise dos factores comportamentais inerentes à pesquisa de informação e gestão e partilha de conteúdos, no âmbito de um estudo em curso sobre estratégias de classificação de informação em instituições de I&D.

A sua colaboração na resposta a este questionário é imprescindível ao presente estudo. Agradecemos a sua contribuição.

Este site observa toda a legislação em vigor sobre protecção de dados, não recolhendo quaisquer elementos sem o prévio conhecimento e autorização dos seus titulares.

A informação recolhida será confidencial e será apenas utilizada para o desenvolvimento do estudo em curso.

Obrigado,

Bruno Eduardo Paulino

Aluno do Mestrado em Gestão de Informação da Faculdade de Engenharia da Universidade do Porto (FEUP) e estagiário do INESC Porto

## 2 DADOS PESSOAIS

### 2.1 Dados profissionais

Instituição:

Departamento:

Função:

Experiência profissional nesta área (anos):

### 2.2 Formação académica

Doutoramento  Mestrado  Licenciatura  Outros

## 3. ORGANIZAÇÃO E PARTILHA DE INFORMAÇÃO

**3.1 Indique se conhece ou já ouviu falar sobre (selecione todas as respostas que se aplicam):**

- ESQUEMA DE CLASSIFICAÇÃO:** organização de documentos numa sequência lógica por assuntos (Encyclopædia Britannica, 2008). Agrupam recursos relacionados numa estrutura hierárquica ou em árvore (Priscilla Caplan, 2003).
- TAXONOMIA:** ciência da localização, descrição, classificação e nomeação de organis-

mos ou coisas. Estes são organizados numa estrutura hierárquica (Wikitionary, 2008).

- FOLKSONOMIA: prática ou método de criação e gestão colaborativa de tags (etiquetas) para marcar e categorizar conteúdos (Wikipédia, 2008).
- ONTOLOGIA: especificação explícita de uma conceptualização, onde conceptualização é uma visão simplificada e abstracta do mundo que desejamos representar para algum propósito. Quando o conhecimento de um dado domínio é representado formalmente, gera um conjunto de objectos e os relacionamentos entre eles (Thomas Gruber, 1993).

### **3.2 Quer individualmente, quer colectivamente, como ORGANIZA/CLASSIFICA/GUARDA documentos e conteúdos (selecione todas as respostas que se aplicam)?**

- Não organizo nem classifico documentos e conteúdos por falta de tempo
- Não organizo nem classifico documentos e conteúdos porque não necessito
- Uso pastas no próprio computador para organizar, classificar e guardar documentos e conteúdos (ex.: por assuntos, ordem alfabética, numérica, cronológica etc.)
- Uso a intranet (ou semelhante) da minha organização e as suas funcionalidades de organização e classificação
- Uso ferramentas de apoio à classificação dos documentos (ex.: thesaurus, cabeçalhos de assunto, esquemas de classificação etc.)
- Não me preocupo muito com isso porque uso um motor de pesquisa no meu computador pessoal ou intranet (ex.: Google Desktop, Windows Sidebar, Spotlight, sistema de pesquisa da própria plataforma, etc.)
- Uso uma aplicação própria para a gestão de referências bibliográficas e documentos electrónicos (tipo EndNote, JabRef, Papers)
- Uso Bookmarking (ex.: favoritos do Internet Explorer, Mozilla Firefox, favoritos da barra de ferramentas do Google, barra de ferramentas do Yahoo etc.)
- Social bookmarking (ex.: CiteUlike, del.icio.us, Flickr, Furl, ma.gnolia, Technorati, Yahoo My Web etc.)
- Uso Wikis (ex.: TWiki, TikiWiki, Wikia, PBWiki, Wikispaces, JotSpot etc.)
- Uso as funcionalidades do gestor do meu Blog (ex: Typepad, Blogspot)
- Uso discos virtuais (ex.: Sky Drive da Microsoft etc.)
- Uso ferramentas de produtividade colaborativas (ex.: Google Docs & Spreadsheets, Zoho Writer, ajaxWrite etc.)

Se utiliza outra forma de organização/classificação, por favor especifique:

---

**3.3 Costuma partilhar, por via electrónica, documentos e conteúdos com (selecione todas as respostas que se aplicam):**

- Colegas do mesmo departamento/grupo de investigação
- Colegas da mesma instituição mas de departamentos/grupos de investigação diferentes
- Colegas de outras instituições
- Não costumo partilhar documentos e conteúdos

**4 FONTES DE INFORMAÇÃO**

**4.1 Para obter a informação que necessita para o seu trabalho de investigação e desenvolvimento, indique com que FREQUÊNCIA utiliza as seguintes fontes (selecione todas as respostas que se aplicam):**

	Sempre	Frequen- temente	As Vezes	Quase nunca	Nunca
Colegas do mesmo departamen- to/grupo de investigação	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Colegas de outros departamen- tos/grupos de investigação	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Colegas externos à instituição	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Biblioteca pessoal (ex.: artigos, livros e outras publicações profis- sionais)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Biblioteca da instituição ou intranet (excepto a biblioteca pessoal)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Biblioteca externa à instituição	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Fontes na internet pesquisadas de uma forma organizada (ex: Meta- Lib, B-on, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Fontes na internet pesquisadas de uma forma não organizada (ex: Google, Google Scholar, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Outros locais específicos na inter- net (ex.: notícias online, grupos de discussão, blogs, sites de projectos etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Publicações externas (ex.: artigos, periódicos científicos etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Bases de dados electrónicas especí- ficas (ex.: Compendex, Inspec, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Se Outra, por favor especifique:					

**4.2 Na escolha e acesso às fontes de informação electrónicas, indique o grau de INFLUÊNCIA dos seguintes factores (selecione todas as respostas que se aplicam):**

	Muito importante	Importante	Indiferente	Pouco importante	Sem importância
Boa qualidade técnica	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Adequação às tarefas	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Actualização	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Relevância	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Fiabilidade	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Qualidade do motor de pesquisa	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Qualidade na classificação de assuntos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Acessibilidade	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Disponibilidade	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Usabilidade	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Experiência com a fonte	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Baixo custo financeiro	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Língua	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Se Outra, por favor especifique:					



**4.3 Para o desenvolvimento do seu trabalho de investigação e desenvolvimento, indique o grau de incidência das DIFICULDADES encontradas na pesquisa de informação (selecione todas as respostas que se aplicam):**

	Sempre	Frequen- temente	As Vezes	Quase nunca	Nunca
Localização da informação	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Excesso de informação	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Informação indisponível	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Informação indesejada	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Utilização do motor de pesquisa	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Interpretação dos resultados obti- dos pelo motor de pesquisa	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Se Outra, por favor especifique:					