

Mauro Pereira Tavares

**CLASSIFICAÇÃO EM GESTÃO DE CONTEÚDOS
EMPRESARIAIS (ECM):**

proposta de método ágil para o desenvolvimento e gestão
de classificações à medida de contextos específicos

Dissertação realizada no âmbito do Mestrado em Ciência da Informação,
orientada pelo Professor Doutor António Manuel Lucas Soares

Faculdade de Engenharia e Faculdade de Letras

Universidade do Porto

Julho de 2013

CLASSIFICAÇÃO EM GESTÃO DE CONTEÚDOS EMPRESARIAIS (ECM):

proposta de método ágil para o desenvolvimento e gestão de classificações à medida de contextos específicos

Mauro Pereira Tavares

Dissertação realizada no âmbito do Mestrado em Ciência da Informação,
orientada pelo Professor Doutor António Manuel Lucas Soares

Membros do Júri

Presidente: Professora Doutora Maria Cristina de Carvalho Alves Ribeiro
(Professora Auxiliar da Faculdade de Engenharia da Universidade do Porto);

Arguente: Professor Doutor Henrique José da Rocha O'Neill (Professor
Associado do ISCTE - Instituto Universitário de Lisboa);

Orientador: Professor Doutor António Manuel Lucas Soares (Professor
Associado da Faculdade de Engenharia da Universidade do Porto);

Julho de 2013

Agradecimentos

Na elaboração deste trabalho foram vários os que contribuíram para que fosse possível atingir o fim desta dissertação.

Em primeiro lugar, cabe-me agradecer ao meu orientador, o Professor Doutor António Lucas Soares, por toda a ajuda, motivação e conselhos dados durante este ano letivo. Agradeço também ao INESC Porto e aos seus colaboradores, por me terem aceitado na sua instituição e colaborado no meu trabalho sempre que o precisei.

Um agradecimento especial à Maria Almeida, colega de mestrado e também colaboradora do INESC Porto, pela ajuda dada na minha integração na instituição, pela sua valiosa colaboração, e por todo o interesse demonstrado no meu trabalho.

Aos meus colegas de mestrado e amigos, Ana, Patrícia, João, Suzi e Cristiana pela companhia, diversão, e disponibilidade ao longo de mais um ano de “escola”.

Um grande obrigado à minha tia, que sempre me apoia e que me retira muitas das responsabilidades domésticas dos meus ombros, o que evitou muito cansaço e dores de cabeça ao longo deste ano.

Aos meus pais, que apesar de estarem relativamente longe, estão sempre disponíveis para me ouvirem e para me ajudarem.

Por último, um agradecimento aos meus quatro miúdos, Bicas, Sassy, Titina, e Tutis, três gatos e uma tartaruga que nunca deixaram de me fazer companhia quando estava mais sozinho, e que nunca deixaram que eu ficasse demasiado imerso no trabalho.

Muito obrigado a todos.

Resumo

O presente trabalho teve como objetivo principal a proposta de um método para o desenvolvimento e gestão de classificações à medida de contextos específicos e contém contribuições para a lacuna existente na literatura científica acerca desta temática.

O método tem como suporte uma perspetiva científica que defende a influência da prática profissional na mudança do paradigma teórico da classificação, onde esta se mostra como um instrumento de apoio às atividades e processos dentro de um contexto organizacional com características específicas, e onde os princípios técnicos da classificação serão úteis como suporte e não como regra. O trabalho foi realizado de acordo com a metodologia de design science, onde se procura resolver determinado problema através da construção e implementação de um novo artefacto. O estudo do problema em contexto profissional permitiu compreender certos padrões de classificação de conteúdos e associar os mesmos a necessidades de recuperação de informação emergentes no contexto. Além disso, deu-se conta da importância das atividades colaborativas e de que forma estas podem ser enquadradas num processo de desenvolvimento e gestão de uma classificação.

Foi construído um método centrado na adequação de modelos estruturais de classificação a determinadas necessidades de recuperação de informação. A avaliação da proposta permitiu constatar o potencial do método como suporte aos profissionais no desenvolvimento e gestão de classificações, no entanto as restrições de tempo não permitiram uma avaliação tão aprofundada quanto desejável.

Concluindo, o trabalho realizado foi suficiente para alcançar os objetivos definidos, e suscita perspetivas de desenvolvimentos futuros, principalmente a avaliação mais aprofundada do método em mais contextos organizacionais, de forma a ser possível monitorizar a utilização das classificações em contexto real de trabalho.

Palavras-chave: classificação; desenvolvimento e gestão de classificações; gestão de conteúdos empresariais; gestão do conhecimento; recuperação de informação; organização de informação; modelo estrutural de classificação; processo colaborativo de gestão de conteúdos.

Abstract

The present work had as main objective the proposal of a method for the development and management of classifications tailored to specific contexts, and contains contributions to the existing gap in the scientific literature about this issue.

The method is supported by a scientific perspective that advocates the influence of professional practice in the change of the theoretical paradigm of classification, where classification is an instrument of support to the activities and processes within a specific organizational context, and where the technical principles of the classification will be useful as a support and not as a rule. The work was carried out in accordance with the methodology of design science, which seeks to solve a given problem by building and implementing a new artifact. The study of the problem in a professional context allowed the understanding of certain patterns of content classification, and associate these with some information retrieval needs emerging in the context. In addition, it was possible to verify the importance of collaborative activities and how these can be included in a process of development and management of a classification.

It was built a method focused in the adequacy of structural models of classification to designated information retrieval needs. The evaluation of the proposal revealed the potential of the method as a support to professionals in the development and management of classification, however the time constraints did not allow an evaluation as thorough as desirable.

In conclusion, the work carried out was enough to reach the defined goals, and raises perspectives of future developments, especially an in-depth evaluation of the method in more organizational contexts, in order to monitor the use of the classifications in real work contexts.

Keywords: classification; development and management of classifications; enterprise content management; knowledge management; information retrieval; information organization; structural model of classification; collaborative content management process.

Lista de ilustrações

Ilustração 1 - Ciclo da gestão de conteúdos, com cinco fases nucleares que podem envolver várias componentes tecnológicas e práticas (assinaladas com cores) _____	35
Ilustração 2 - Comparação das fases do ciclo de gestão de conteúdos (Kampffmeyer, 2005; AIIM, 2012) com o ciclo de gestão da informação (Pinto e Silva, 2005) _____	39
Ilustração 3 - Tag cloud _____	57
Ilustração 4 - Exemplo de mapa de conceitos _____	61
Ilustração 5 - Fases do processo de design science _____	85
Ilustração 6 - Estrutura de pastas em hierarquia utilizada no Plone _____	93
Ilustração 7 - Classificação hierárquica utilizada na Dropbox utilizada na área de consultoria do INESC Porto, adequada ao sistema através de uma estrutura hierárquica de pastas organizada segundo as fases de gestão de um projeto _____	94
Ilustração 8 - Funcionalidade de atribuição de categorias do Plone, um dos possíveis métodos para "classificar" o conteúdo da plataforma _____	101
Ilustração 9 - Funcionalidade de atribuição de etiquetas a conteúdos do Windows Explorer _____	102
Ilustração 10 - Representação gráfica da proposta de método para o desenvolvimento e gestão de classificações à medida _____	121

Lista de tabelas

Tabela 1 - Vários tipos de ferramentas para a gestão de conteúdos empresariais _____ 74

Tabela 2 - Alguns dos softwares mais populares da área de Enterprise Search _____ 79

Lista de abreviaturas e siglas

BI – Business Intelligence;

CI – Ciência da Informação;

ECM – Enterprise Content Management;

IA – Inteligência Artificial;

KOP – Knowledge organization problem.

Sumário do conteúdo

0. INTRODUÇÃO.....	21
0.1 CONTEXTUALIZAÇÃO DO TRABALHO	21
0.2 OBJETIVOS E RESULTADOS ESPERADOS	23
0.2.1 LISTA DE OBJETIVOS	23
0.2.2 RESULTADOS ESPERADOS:.....	24
0.3 CONTRIBUIÇÕES DO TRABALHO	25
0.4 ESTRUTURA DA DISSERTAÇÃO	26
1. REVISÃO DE LITERATURA.....	27
1.1 REVISÃO DA LITERATURA CIENTÍFICA	27
1.1.1 ECM - ENTERPRISE CONTENT MANAGEMENT	28
1.1.2 CICLO DE GESTÃO DE CONTEÚDOS	34
1.1.3 CLASSIFICAÇÃO	40
1.1.3.1 O conceito “Classificação”	42
1.1.3.2 Classificação, categorização e indexação	43
1.1.3.3 Princípios da classificação	47
1.1.3.4 Modelos estruturais de sistemas de classificação	52
1.1.3.5 Classificação para além da visão tradicional da Ciência da Informação	61
1.1.3.6 Influência da prática profissional na evolução teórica do conceito.....	63
1.1.4 PERSPETIVA DEFENDIDA NESTE TRABALHO E PRINCIPAIS IDEIAS A RETER	67
1.2 ESTADO-DE-ARTE TECNOLÓGICO	70
1.2.1 ESTRATÉGIAS DE ECM	70
1.2.2 TECNOLOGIAS DE ECM	72
1.2.3 BUSINESS INTELLIGENCE E ENTERPRISE INFORMATION MANAGEMENT	76
1.2.4 ENTERPRISE SEARCH	78
1.2.5 ENTERPRISE 2.0	80
1.2.6 IDEIAS A RETER QUANTO AO ESTADO-DE-ARTE DA GESTÃO DE CONTEÚDOS	81
2. METODOLOGIA DE INVESTIGAÇÃO	83
2.1 DESIGN SCIENCE	83
2.2 ADAPTAÇÃO DA METODOLOGIA AOS OBJETIVOS DA DISSERTAÇÃO	86

3. PADRÕES DE ORGANIZAÇÃO E CLASSIFICAÇÃO DE CONTEÚDOS ...	90
3.1 ESTRUTURAÇÃO DA CLASSIFICAÇÃO	92
3.2 ADAPTAÇÃO DA CLASSIFICAÇÃO AO SISTEMA DE GESTÃO DE CONTEÚDOS	100
3.3 DEFINIÇÃO DOS CONCEITOS, TERMOS E RELAÇÕES	104
 4. PROCESSOS COLABORATIVOS PARA O DESENVOLVIMENTO E GESTÃO DE CLASSIFICAÇÕES	 108
4.1 DESENVOLVIMENTO COLABORATIVO DA CLASSIFICAÇÃO	109
4.2 UTILIZAÇÃO E GESTÃO COLABORATIVA DA CLASSIFICAÇÃO	112
4.3 PLANEAMENTO DO PROCESSO DE DESENVOLVIMENTO E GESTÃO DA CLASSIFICAÇÃO	115
 5. MÉTODO PARA O DESENVOLVIMENTO E GESTÃO DE CLASSIFICAÇÕES À MEDIDA DE CONTEXTOS ESPECÍFICOS	 118
5.1 DESCRIÇÃO DO MÉTODO	120
5.1.1 MÉTODO: PRINCÍPIOS	120
5.1.2 MÉTODO: PROCESSO	121
5.1.3 MÉTODO: INDICAÇÕES INSTRUTÓRIAS	124
5.2 AVALIAÇÃO DA APLICAÇÃO PRÁTICA DO MÉTODO	130
 6. CONCLUSÕES E PERSPETIVAS DE DESENVOLVIMENTO	 139
 REFERÊNCIAS BIBLIOGRÁFICAS	 146
 ANEXOS	 157
ANEXO 1 - GUIÃO SIMPLIFICADO DE APLICAÇÃO DO MÉTODO PARA DESENVOLVIMENTO E GESTÃO DE CLASSIFICAÇÕES À MEDIDA	157
ANEXO 2 - VERSÃO PROVISÓRIA E RESUMIDA DO MÉTODO, UTILIZADA NA AVALIAÇÃO	160
ANEXO 3 - FORMULÁRIO PARA O REGISTO DE INFORMAÇÃO NA AVALIAÇÃO	163
ANEXO 4 - TRANSCRIÇÃO DAS ENTREVISTAS REALIZADAS	169
ENTREVISTA A “ANÓNIMO 1”	169
ENTREVISTA A “ANÓNIMO 2”	174
ENTREVISTA A “ANÓNIMO 3”	178

o. Introdução

o.1 Contextualização do trabalho

A organização e classificação da informação são aspetos fundamentais da gestão da informação. No mundo empresarial, essa importância é evidente, especialmente no que diz respeito a gestão dos conteúdos empresariais que são criados e recebidos diariamente na atividade de uma organização.

As organizações confrontam-se cada vez mais com uma grande quantidade e diversidade de conteúdos necessários aos seus processos organizacionais, que necessitam de ser geridos da melhor forma possível. Esta gestão exige métodos, técnicas e práticas de organização e classificação da informação que permitam que as organizações recuperem esses conteúdos com elevada eficácia e eficiência. A quantidade de informação criada e armazenada em suportes físicos e, especialmente, em suportes eletrónicos tem crescido exponencialmente – alguns estudos apontam para um crescimento de cerca de 800MB de conteúdos por pessoa por ano (Gingel, 2006, citado em Alalwan e Weisroffer, 2012), o que em organizações com maiores dimensões se poderá revelar como um problema em pouco tempo, se não forem pensadas soluções atempadamente.

Os sistemas de gestão de conteúdos empresariais têm sido a principal aposta da indústria para lidar com o aumento acelerado da quantidade de informação produzida. O termo *Enterprise Content Management (ECM)* foi cunhado por uma das maiores associações internacionais envolvidas nesta área, a AIIM (Association for Information and Image Management), sendo agora universalmente utilizado, tanto no mundo dos negócios, como no da investigação. No entanto, as ferramentas e aplicações tecnológicas não são, por si só, suficientes para uma organização conseguir gerir eficazmente os seus conteúdos – o contexto e as características específicas de cada organização exigem que seja feita a personalização dos sistemas tecnológicos para que se atinja o melhor alinhamento possível entre a tecnologia e o contexto organizacional. E é neste aspeto que muitos projetos de ECM falham: apesar da tecnologia ter as

funcionalidades necessárias, a organização não as consegue personalizar e adaptar aos requisitos da organização.

Grande parte da adaptação envolve o desenvolvimento e gestão de classificações para os conteúdos: é a classificação dos conteúdos criados que permite a sua organização e a sua recuperação, que obviamente trazem melhorias em termos de eficiência organizacional (informação recuperada mais rapidamente significa menos tempo por tarefa) e em termos de otimização dos processos e atividades organizacionais (informação “perdida” pode representar enormes riscos). E são estas melhorias de eficácia e eficiência as principais preocupações da indústria de ECM, o que comprova a importância e a necessidade de apostar no desenvolvimento de melhores métodos para a gestão de conteúdos e, por consequência, para a classificação eficaz dos mesmos (AIIM, 2011).

No entanto, é difícil estabelecer métodos de organização, gestão e classificação de conteúdos universais, isto é, que se ajustem a todos os contextos - há que considerar as características específicas de cada organização e as plataformas que utilizam para gerir a informação, o que leva a que exista a necessidade cada vez mais urgente de métodos e técnicas adaptadas a cada caso. Além disso, o facto de grande parte da informação criada numa organização ser não-estruturada e não estar apresentada em formatos standardizados, dificulta o processo de classificação e recuperação dos conteúdos. As estimativas indicam que 80% da informação de uma organização não está estruturada, e que esta cresce a um ritmo entre 65% a 200% por ano (EMC Corporation, 2006, citado em Alalwan e Weisroffer, 2012), o que torna o desenvolvimento de melhores métodos de classificação e melhores plataformas de gestão de conteúdos uma questão emergente e essencial para a gestão da informação organizacional – e por, consequência, para aumentar a eficiência e eficácia das atividades e processos organizacionais.

Nesta perspetiva, surge então a oportunidade de ajudar a prática profissional no desenvolvimento e gestão de sistemas de classificação, através de um método suficientemente flexível para se adaptar a diferentes contextos. Este trabalho relata, assim, os resultados do processo de investigação realizado na sua construção.

o.2 Objetivos e resultados esperados

De seguida, está apresentada a árvore e a lista dos objetivos definidos para este trabalho, sendo seguida por uma descrição mais aprofundada dos resultados esperados para cada um dos mesmos:

o.2.1 Lista de objetivos

1. Desenvolver e avaliar um método eficaz para o desenvolvimento e gestão de classificações à medida de contextos específicos

1.1. Definir uma visão que guie a criação de classificações à medida de forma eficaz

- 1.1.1. Estabelecer pontos de contacto e de divergência na discussão existente dentro da área da Ciência da Informação acerca da visão do sistema conceptual centrado nos conceitos de organização, classificação e recuperação de informação
- 1.1.2. Caracterizar a visão da Inteligência Artificial, da indústria de ECM e das associações especializadas em ECM no que diz respeito à classificação
- 1.1.3. Caracterizar o papel do profissional de informação na gestão e classificação da informação

1.2. Identificar padrões para a organização e classificação de conteúdos

- 1.2.1. Identificar e caracterizar padrões de organização e classificação de conteúdos para vários tipos de contextos empresariais
- 1.2.2. Identificar e caracterizar tipos de plataformas de gestão de conteúdos para vários tipos de contextos empresariais

1.3. Especificar processos colaborativos para o desenvolvimento e gestão da meta-informação e da classificação

- 1.3.1. Identificar e caracterizar exemplos de atividades colaborativas no desenvolvimento de classificações envolvendo múltiplos colaboradores;
- 1.3.2. Identificar e caracterizar plataformas e ferramentas adequadas para a gestão colaborativa da meta-informação e da classificação
- 1.3.3. Caracterização do papel dos colaboradores e da organização na gestão colaborativa da informação

0.2.2 Resultados esperados:

Sendo o trabalho de natureza exploratória, grande parte dos objetivos passam por analisar criticamente, interpretar e relacionar informação das mais variadas fontes, de forma a ser possível desenvolver uma visão fundamentada e uma base de conhecimento sólida para guiar o processo de criação do método de construção de classificações à medida – que é o objetivo principal da dissertação. Por este motivo, é um pouco complicado definir objetivos intermédios que pressuponham resultados tangíveis, uma vez que o trabalho a realizar será sobretudo de análise e interpretação de informação, para suportar a construção do método (objetivo principal); os resultados deste tipo de objetivos são apenas comprováveis através da redação escrita, que será representativa do pensamento interpretativo.

Posto isto, o resultado principal que se pretende obter é um método eficaz para o desenvolvimento e gestão de classificações à medida. Espera-se assim um método adaptável a vários contextos, eficaz, capaz de ser gerido colaborativamente por vários colaboradores, de aplicação relativamente simples e orientado para a prática (com indicações instrutórias).

Tendo em conta a natureza do trabalho, os resultados tangíveis para os objetivos intermédios esperados passam, então, por uma síntese da informação analisada e interpretada, informação que deverá ser suficiente para corresponder às exigências de cada um dos três objetivos intermédios (1.1, 1.2 e 1.3).

Assim no **objetivo 1.1**, pretende-se explicar, de forma fundamentada, a necessidade de uma visão única para apoiar a construção de classificações, através da discussão das visões existentes dentro da área da Ciência da Informação acerca do sistema conceptual centrado nos conceitos de organização, classificação e recuperação de informação. Se possível, pretende-se ainda, neste objetivo, analisar também a visão da área da Inteligência Artificial e da indústria de ECM em relação a estas questões mais concetuais e teóricas, bem como caracterizar o papel do profissional da informação nesta área.

No que diz respeito ao **objetivo 1.2.**, espera-se identificar e caracterizar padrões de organização e classificação de conteúdos e tipos de ferramentas e plataformas de

gestão de conteúdos para várias áreas empresariais, com a finalidade de conseguir que o método de desenvolvimento de classificações à medida a construir possa ter o máximo de adaptabilidade possível (que é um dos grandes objetivos do método que se pretende criar).

Por último, no **objetivo 1.3.** pretende-se identificar, caracterizar e analisar processos colaborativos de desenvolvimento e gestão de classificações, para que possam ser adaptadas ao método a construir, assim como o próprio papel e esforço necessário do colaborador nestas atividades.

0.3 Contribuições do trabalho

Tendo em conta a informação apresentada, podemos verificar que este trabalho de dissertação traz contribuições em várias frentes: em primeiro lugar, discutirá aspetos mais concetuais e teóricos relacionados com a classificação de conteúdos, contribuindo assim para o preenchimento da lacuna existente na literatura nessa dimensão; em segundo lugar, o objetivo desta dissertação é a criação de um método ágil para o desenvolvimento e gestão de classificações à medida, o que vem corresponder à falta de literatura relacionado com a dimensão processual, ou seja, com a sistematização do processo de implementação de sistemas ECM; em terceiro lugar, este trabalho é suportado pela perspetiva da ciência da informação, e, como tal, não incidirá em demasia na dimensão tecnológica de ECM, mas sim numa conjugação entre a teoria, a tecnologia, e a prática profissional, o que não tem sido explorado na publicação científica; por último, e talvez mais importante, os resultados deste trabalho pretendem ser um auxílio (ainda que careça investigação posterior) para os profissionais e para as organizações que necessitam de mais guidelines e orientações práticas no desenvolvimento e gestão das classificações, e na própria adaptação dos sistemas aos contextos específicos.

0.4 Estrutura da dissertação

A dissertação está estruturada em capítulos, sendo que cada um destes contém uma série de secções e, por vezes, subsecções. No final, são ainda incluídos alguns anexos com informação acessória ao corpo central do trabalho. Em termos geral, os capítulos tratam dos seguintes temas:

- Cap. 0** - Introdução ao trabalho, onde é feita a contextualização do tema e apresentados os objetivos e principais contribuições da investigação;
- Cap. 1** - Revisão de literatura científica das áreas científicas relacionadas com a gestão de conteúdos empresariais, bem com do estado-de-arte da indústria de ECM;
- Cap. 2** - Metodologia de investigação em que se enquadrrou o trabalho realizado nesta dissertação;
- Cap. 3** - Análise da informação obtida em contexto profissional acerca de padrões de organização e classificação de conteúdos, bem como as conclusões retiradas em relação a este aspeto;
- Cap. 4** - Análise da informação obtida em contexto profissional acerca de processo colaborativos para o desenvolvimento e gestão de classificação, bem como as conclusões retiradas em relação a este aspeto;
- Cap. 5** - Apresentação, descrição e avaliação da proposta de método para o desenvolvimento e gestão de classificações à medida de contextos específicos;
- Cap. 6** - Conclusões e considerações finais da dissertação e perspetivas de trabalho futuro.

1. Revisão de literatura

A gestão de conteúdos empresariais é uma área inserida num âmbito complexo, onde não só estão envolvidos aspetos conceituais, técnicos e práticos ligados à gestão da informação, mas também elementos de cariz mais tecnológico ligados às áreas informáticas. Tendo em conta que os objetivos deste trabalho têm componentes teóricas e conceituais, mas numa perspetiva de suporte a uma área fundamentalmente prática e profissional (ECM), é importante que a revisão de literatura incida não só na discussão conceitual do sistema conceitual ligado à classificação, gestão e recuperação de informação, mas também na parte tecnológica e prática, ligada à indústria de ECM e às aplicações tecnológicas desenvolvidas no contexto desta.

Neste sentido, esta revisão de literatura está dividida em duas grandes secções, em que a primeira incide na revisão da literatura científica publicada relacionada com a temática do trabalho, e a segunda procura discutir aspetos ligados ao estado de arte tecnológico da área de gestão de conteúdos empresariais.

1.1 Revisão da literatura científica

A classificação, temática central deste trabalho, é um dos elementos mais importantes na gestão e organização da informação, e por conseguinte na Ciência da Informação. Neste trabalho, a classificação é discutida e investigada numa perspetiva de apoio teórico e prático à gestão de conteúdos empresariais (ECM), e, dessa forma é importante perceber não só a natureza, os objetivos e as características da ECM, mas também qual a relação destes aspetos com a classificação. Além disso, o sistema conceitual da classificação é alvo de bastante discussão na literatura, visto de distintas formas em diferentes comunidades científicas, e verificando-se mesmo perspetivas divergentes dentro da área da Ciência da Informação. Os objetivos principais desta revisão são, então, relacionar a ECM com a classificação, estudar e comparar as visões divergentes acerca do conceito de classificação, e problematizar toda esta discussão de forma a encontrar e definir a perspetiva mais adequada a suportar a construção do método no âmbito da área de ECM.

Nesta perspetiva, a discussão das próximas páginas começará por caracterizar os aspetos ligados à área da gestão de conteúdos empresariais, especialmente o que diz respeito ao ciclo de gestão de conteúdos, de forma a ser possível relacionar a atividade geral de classificação com a área prática de ECM. Estabelecida esta relação, será possível depois passar à problematização da discussão existente acerca do sistema concetual da classificação, dando ênfase à divergência de opiniões na área da Ciência da Informação, à perspetiva da Inteligência Artificial, e à relação entre classificação, categorização e indexação. Não obstante, serão também analisadas questões mais práticas ligadas à classificação, tais como os princípios da classificação e os modelos estruturais, que serão também essenciais no apoio à construção do método. No final desta secção, pretende-se chegar à definição de uma perspetiva teórica que suporte o método.

1.1.1 ECM - Enterprise Content Management

A AIIM considera que ECM engloba “as estratégias, métodos e ferramentas usadas para capturar, gerir, armazenar, preservar e transmitir conteúdo e documentos relacionados com os processos organizacionais”, sendo que “as ferramentas e estratégias de ECM permitem a gestão da informação não estruturada, onde quer que ela se encontre” ¹. A partir desta definição conseguimos detetar, pelo menos, três aspetos essenciais e caracterizadores da gestão de conteúdos empresariais – então, segundo a AIIM, a **ECM envolve**:

- Estratégias, métodos e ferramentas;
- Um ciclo de gestão dos conteúdos, desde a sua captura até à sua transmissão;
- Gestão de informação não-estruturada (ou seja, dos conteúdos).

Apesar da AIIM ocupar uma posição importante dentro da área – foi mesmo a responsável pela criação do termo *Enterprise Content Management* -, a sua definição não é adotada de modo consensual por todos os investigadores e profissionais da área, existindo assim várias pontos de vista no que diz respeito à verdadeira essência da

¹ AIIM (2011). ‘State of the ECM Industry: how well is it meeting business needs’. AIIM, acedido a 5 de dezembro de 2012, <http://www.aiim.org/Research-and-Publications/Research/Industry-Watch/State-of-the-ECM-Industry-2011>

ECM. Por exemplo, Smith e McKeen (citado em Alalwan e Weisroffer, 2012) levam a definição mais além, ao considerar que a área engloba todas as estratégias, ferramentas, técnicas e processos necessários para a gestão de todos os ativos informacionais, independentemente do seu tipo – por outras palavras, Smith e McKeen consideram que a ECM envolve não só a gestão da informação não-estruturada, mas também de todos os restantes tipos de informação organizacional. Outros autores seguem esta linha de pensamento (de Smith e McKeen), propondo definições para a área que se aproximam da “globalidade” da definição proposta pelos dois canadianos: Boiko (citado por Alalwan e Weisroffer, 2012), por exemplo, vê a ECM como uma “evolução da gestão da informação que envolve a gestão de conteúdo estruturado e não-estruturado durante todo o ciclo de vida dos conteúdos”². Contudo, parece-me que esta visão é bastante próxima (se não idêntica) à definição de gestão de informação, uma vez que não fazem distinção entre tipos de informação, não estabelecem diferenças entre o que entendem por conteúdo e por informação e, desse modo, acabam por encarar a ECM como algo que envolve a gestão de toda a informação organizacional durante todo o seu ciclo de vida, bem como as estratégias, métodos e tecnologias utilizadas para essa finalidade – e este ponto de vista parece-me confundir-se com um conceito (e área) já bem definido e estudado, designado por gestão da informação (Deltci, 2008; Pinto e Silva, 2005).

Tendo em conta esta ideia, é essencial que os autores definam as fronteiras da ECM, de forma a evitar confusões e até mesmo para justificar a existência da própria área (“qual o motivo de definir uma nova área se a sua finalidade é a mesma de uma já existente?”). Neste aspeto, penso que a AIIM, bem como vários outros autores (Gartner, 2012; O’Callaghan e Smits, 2005; Munkvold et al, citado em Alalwan e Weisroffer, 2012), são eficazes, uma vez que as suas propostas conseguem restringir a ECM à gestão da informação não-estruturada e, desse modo, definem uma área de ação mais concreta. **O facto de limitarem a ECM a um tipo de informação ajuda também a relacionar a área com a gestão da informação – se a ECM envolve apenas a gestão de informação não-estruturada ou conteúdos, é lógico pensar que estará enquadrada dentro da área da gestão da**

² B. Boiko, citado em Alalwan, J. e H. Weisroffer H. (2002). “Enterprise content management research: a comprehensive review”, *Journal of Enterprise Information Management* **25**(5), 442, acessado a 20 de novembro de 2012, <http://www.emeraldinsight.com/journals.htm?articleid=17054879>.

informação. Esta visão parece ser claramente mais realista e lógica, aproximando-se, de certo modo, da visão da Ciência da Informação no que diz respeito à gestão da informação em organizações (Pinto e Silva, 2005), **que coloca a gestão de conteúdos como uma parte essencial do ciclo de gestão de informação** (que veremos mais à frente quando discutirmos as possíveis relações entre o ciclo de gestão de conteúdos e o ciclo de gestão de informação). Contudo, o facto da “estrutura” da informação ser algo que suscita sempre dúvidas (qual é o ponto fulcral em que deixamos de ter informação não-estruturada e passamos a ter informação estruturada?) torna difícil a definição do foco específico da ECM. Assim, e para efeitos deste trabalho, considera-se informação estruturada como aquela que é passível de ser facilmente capturada pelos sistemas de informação, uma vez que possui uma estrutura normalizada (como por exemplo formulários, faturas, recibos e toda a documentação criada em grande quantidade e com formato e estrutura idêntica e normalizada).

De referir ainda que um grupo de autores considera a ECM como um tipo ou uma sub-área da Gestão do Conhecimento (KM) (Duffy, 2001; Carvalho e Ferreira, citado em Alalwan e Weistroffer, 2001; Päivärinta e Monkvoid, 2005). Apesar desta linha de pensamento parecer diferente da acima descrita, é preciso notar que grande parte dos artigos que referem esta ideia foram publicados durante o *boom* da literatura sobre gestão do conhecimento e, tal como muitos artigos publicados nesse período, sofrem de uma duvidosa aplicação do termo “gestão do conhecimento” – isto é, aplicam este termo ao conceito de gestão da informação, de forma a “promover” a publicação como algo novo. Relativamente a este aspeto, a opinião de Tom Wilson³ parece-me a mais a mais realista, uma vez que a “gestão do conhecimento” é, no fundo, uma evolução natural da gestão da informação (e não uma revolução que justifique a criação de um novo conceito). Desta forma, penso que o grupo de autores que segue esta linha de pensamento acaba por apoiar as ideias apresentadas no último parágrafo, dado que coloca a ECM como uma sub-área da KM, que no fundo é gestão de informação.

Por outro lado, o facto do termo ECM ter sido originado na indústria (AIIM) e só depois ter passado para o lado da investigação científica, explica grande parte das dificuldades na definição epistemológica do próprio conceito e da delimitação da área:

³ Ver Wilson, T. (2002). “The nonsense of 'knowledge management'”, *Information Research* 8(1), acedido a 7 de janeiro de 2013, <http://informationr.net/ir/8-1/paper144.html>

a origem de “novas áreas” e novos termos na indústria é (quase) sempre provocada por novos avanços tecnológicos (ao contrário da investigação, onde o surgimento de novos conceitos necessita de ser fundamentado epistemologicamente), que precisam de ser alvos de marketing para terem maior impacto – **a definição de um novo termo para um grupo de novas tecnologias é normalmente uma forma simples e eficaz de dar visibilidade às novas descobertas e de despoletar novas investigações baseadas no supostamente “novo conceito”**. E de certo modo, consegue-se, sem muito esforço, estabelecer uma linha evolutiva de “conceitos” promovidos pela indústria das tecnologias de informação, desde a gestão da informação, passando pela gestão do conhecimento, até à ECM atual – conceitos que, porventura, têm todas características bastantes semelhantes aos olhos da indústria e que, por isso, não são mais que um simples desenvolvimento natural da definição de “gestão da informação” provocado pelas mudanças tecnológicas e sociais.

Posto isto, podemos concluir que **mesmo que a área de ECM inclua por vezes características que fogem a esta última descrição da AIIM, como por exemplo a gestão de documentos e de documentos de arquivo, é seguro afirmar que o núcleo da sua finalidade é a gestão da informação não-estruturada nas organizações**. A complexidade e dificuldade inerentes à gestão de informação não-estruturada exigem estratégias, métodos, ferramentas e técnicas mais avançadas que aquelas utilizadas na gestão de informação com uma estrutura definida. A existência da ECM é também justificada pela importância atual e futura da web para as organizações – na web, grande parte dos websites são totalmente compostos por conteúdos, cuja apresentação ao utilizador necessita de ser gerida e estruturada da melhor forma possível, e para isso os desenvolvimentos da área de ECM assumem um papel essencial, uma vez que introduzem novos avanços tecnológicos para a gestão desse tipo de informação. Por estes motivos, a posição da ECM como uma área separada na investigação e na indústria é logicamente compreensível e aceitável, mesmo que epistemologicamente possa existir discussão quanto às fronteiras do “conceito” – assim, e por causa das razões epistemológicas acima discutidas, **será mais certo e adequado entender a ECM como uma área, do que como um conceito**.

Apesar desta divergência de opiniões quanto aos tipos de informação relevantes para a ECM, parece consensual entre os vários autores que esta área envolve não apenas as tecnologias práticas, mas também outros aspetos organizacionais como estratégias, processos, técnicas e métodos, que fazem da ECM uma área transversal a todo o contexto organizacional e essencial para a eficiência e eficácia dos processos das organizações nesta Sociedade da Informação. O método descrito neste trabalho pretende focar a gestão e classificação da informação não-estruturada

A literatura identifica alguns dos principais impactos da ECM nas organizações, baseando-se em casos de estudo reais, nomeadamente (Päivärinta e Munkvold, 2005):

- ✓ Maior colaboração interna e externa, envolvendo a partilha de informação;
- ✓ Maior valorização dos serviços e mais variedade de serviços através da utilização dos conteúdos;
- ✓ Maior qualidade e fiabilidade na informação;
- ✓ Modernização da imagem da empresa;
- ✓ Maior eficiência, eficácia e flexibilidade na gestão da informação e nos processos organizacionais, através da reutilização de conteúdo;
- ✓ Redução do trabalho rotineiro, maior facilidade nas tarefas;
- ✓ Aumento da memória organizacional;
- ✓ Redução de custos relacionados com o processamento de informação;
- ✓ Satisfação de regulamentos e standards externos acerca da gestão da informação organizacional.

Apesar do reconhecimento da importância da área, e de todas as possíveis mais-valias que pode trazer a uma organização, a literatura técnico-científica relata poucos trabalhos que sistematizem o desenvolvimento de métodos, técnicas e práticas de desenvolvimento de classificações ajustadas às diferentes características, contextos e processos organizacionais, que seriam bastante importantes para auxiliar as organizações no processo de adaptação de um sistema de ECM aos requisitos organizacionais. Especialmente no que diz respeito às dimensões mais conceituais e teóricas da ECM, nos últimos dez anos apenas uma pequena percentagem das publicações abordaram esse aspeto como principal foco do trabalho, ao contrário das

dimensões mais tecnológicas e técnicas que têm sido o motor de grande parte da investigação na área (Alalwan e Weisroffer, 2012). Outro ponto interessante é o facto de também a dimensão processual da investigação em ECM ser pouco estudada – **isto indica que existe uma lacuna na literatura no que diz respeito à sistematização prática e instrutória das tarefas (as típicas *guidelines*), passe o estrangeirismo) para a implementação bem sucedida de um sistema de ECM, bem como uma falta de investigação baseada em dados empíricos relacionados com projetos desta natureza.** Este facto também parece ser uma consequência da escassez de projetos bem sucedidos nesta área. Outro aspeto interessante é facto dos problemas relacionados com a implementação de meta-informação e de adaptação de taxonomias ao sistema são mais retratados na literatura profissional do que na literatura académica (Päivärinta e Monkvoll, 2005), o que mostra, por um lado, a importância das questões de classificação de informação e de representação do conhecimento para as organizações, e, por outro lado, a escassez de literatura científica sobre o desenvolvimento e implementação prática desses dois aspetos.

Além disso, as análises da literatura publicada nesta área revelam que existe uma certa tendência da própria investigação para deixar de parte dos seus frameworks de análise a fase de adaptação e personalização do sistema: isto é, dentro do ciclo de desenvolvimento e implementação de um sistema de gestão de conteúdos, colocam em segundo plano a adaptação do sistema, focando antes a fase de “integração de sistemas”, e assim dando ênfase à interoperabilidade e à integração do novo sistema com os já existentes (Alalwan e Weisroffer, 2012). Isto pode indicar uma tendência da investigação para seguir os padrões da indústria, ou seja, de focar os aspetos relacionados com o desenvolvimento técnico de novas e melhores aplicações, em vez de tentar problematizar e explicar o processo prático de implementação da ECM para ajudar as organizações – mais uma vez, parece existir uma certa tendência da indústria em colocar o foco no desenvolvimento de melhor tecnologia, e de descurar o processo de implementação, o que leva a que as organizações optem pela solução mais imediata de implementar mais e mais aplicações. **Esta tendência origina problemas de interoperabilidade nas organizações, que por sua vez fomentam o surgimento de novas tecnologias e investigações para os tentarem resolver – isto dá origem a um ciclo vicioso que acaba por “mascarar” o epicentro**

dos problemas, ou seja, a implementação e adaptação dos sistemas ECM aos requisitos e contextos organizacionais.

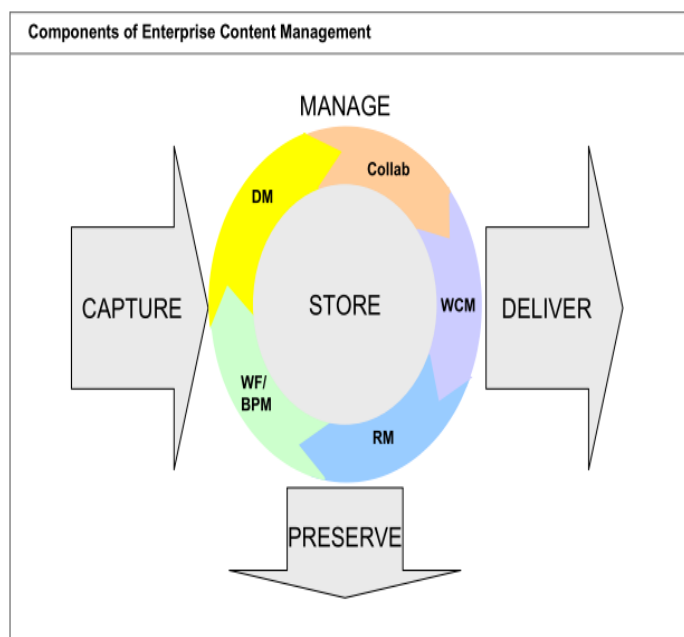
A revisão das questões mais conceituais sobre a área de ECM e do cerne da literatura publicada nos últimos anos permitiu-nos compreender a importância das classificações para a gestão adequada dos conteúdos (isto é, na gestão da informação não-estruturada), e, por consequência, para a área da ECM. A classificação é então essencial para a gestão e recuperação dos conteúdos, ao permitir que se construa uma organização lógica da informação, sempre de acordo com os requisitos específicos de cada contexto. Então, é necessário agora compreender onde “entra” a classificação dentro do ciclo de atividades relacionadas com a gestão de conteúdos, para podermos definir o seu âmbito antes de discutirmos questões mais conceituais e práticas sobre o conceito “classificação”. A revisão deste aspeto permite também uma visão geral das atividades envolvidas neste processo holístico e cíclico de gestão de conteúdos.

1.1.2 Ciclo de gestão de conteúdos

A revisão das questões mais conceituais sobre a área de ECM e do cerne da literatura publicada nos últimos anos permitiu-nos compreender a importância das classificações para a gestão adequada dos conteúdos (isto é, na gestão da informação não-estruturada), e, por consequência, para a área da ECM. A classificação é então essencial para a gestão e recuperação dos conteúdos, ao permitir que se construa uma organização lógica da informação, sempre de acordo com os requisitos específicos de cada contexto. Então, é necessário agora compreender onde “entra” a classificação dentro do ciclo de atividades relacionadas com a gestão de conteúdos, para podermos definir o seu âmbito antes de discutirmos questões mais conceituais e práticas sobre o conceito “classificação”. A revisão deste aspeto permite também uma visão geral das atividades envolvidas neste processo holístico e cíclico de gestão de conteúdos.

O ciclo de gestão de conteúdos, baseado na área da ECM, proposto pela AIIM e relativamente consensual entre os investigadores e os profissionais, apresenta cinco fases nucleares (ver ilustr. 1)⁴:

- ✓ **Capturar** – esta fase envolve a captura, preparação e processamento dos conteúdos organizacionais criados ou obtidos de outras fontes. A preparação e processamento dos conteúdos pode representar a simples digitalização de documentos físicos (através de variadas tecnologias de reconhecimento ótico), até ao processamento de conteúdos eletrónicos, independentemente do suporte. A captura da informação envolve necessariamente a criação de meta-informação para descrição das suas características. Dado ao crescimento exponencial da quantidade de informação organizacional, esta meta-



informação é, na maioria dos casos, criada de forma automática aquando do

Ilustração 1 - Ciclo da gestão de conteúdos, com cinco fases nucleares que podem envolver várias componentes tecnológicas e práticas (assinaladas com cores)

processamento da informação, através de métodos de indexação automática e, também, de classificação automática. Estes dois últimos conceitos são muitas vezes usados como sinónimos na literatura, no entanto são diferentes e devem ser distinguidos (veremos este assunto mais à frente neste capítulo). A indexação e classificação automática da informação realizada nesta fase têm influência nas restantes fases do ciclo de gestão de conteúdos, uma vez que é a

⁴ Ilustração retirada de Kampffmeyer, U. (2004). "Trends in Record, Document and Enterprise Content Management". Project Consult, Hamburgo, acedido a 20 de dezembro de 2012, http://twiki.bes-sa.org/pub/Development/SystemsPlanningInformationSystem/ECM_Handout_english_SER.pdf.

meta-informação criada nessas duas atividades que permite que os conteúdos sejam geridos, armazenados, recuperados e transmitidos de forma eficaz e eficiente – isto significa que os conteúdos incorretamente classificados e/ou indexados poderão ser bastante difíceis de recuperar mais tarde, o que criará problemas nos processos organizacionais e também na partilha de informação organizacional. As classificações assumem assim um papel especialmente importante nesta fase do ciclo (Kampffmeyer, 2004; AIIM, 2012).

✓ **Gerir** – esta fase explica-se por si mesma, uma vez que envolve, realmente, a gestão dos aspetos relacionados com a organização, processamento e utilização da informação. As componentes a colorido na figura 1 são geridas também nesta fase (Kampffmeyer, 2004; AIIM, 2012):

- **Gestão documental (DM)**– envolve os aspetos relacionados com a gestão de documentos com informação não-estruturada. Este aspeto engloba, muitas vezes, algo que vai mais além da gestão de conteúdos, nomeadamente a gestão de qualquer tipo de documento (mais uma vez, percebe-se que grande parte da indústria e da investigação tendem a considerar a ECM como uma evolução direta da gestão da informação, e daí que seja frequente encontrarmos áreas dentro da ECM que deveriam ser áreas independentes, tal como a gestão documental e a gestão de documentos de arquivo).
- **Atividades colaborativas (Collab)**– processos, atividades e tarefas colaborativas que contribuem para a eficácia e eficiência da gestão dos conteúdos organizacionais.
- **Gestão de conteúdos web (WCM)**
- **Gestão de documentos de arquivo (RM)** – esta componente é bastante discutível, uma vez que a informação que é avaliada e considerada importante para preservar, deverá já ter sido alvo de alguma estruturação (e aqui saímos do âmbito da ECM e passamos ao âmbito da gestão do arquivo, que é uma área relacionada mas independente da ECM).
- **Gestão de processos/workflow (WF/BPM)** – gestão e articulação das atividades relacionadas com a gestão dos conteúdos com os processos organizacionais, bem como a gestão do workflow de todos os processos envolvidos no ciclo de gestão de conteúdos empresariais .

Vistas as componentes envolvidas na fase de gestão do ciclo, poderemos concluir a inclusão de algumas componentes – gestão de processos, gestão documental, gestão de documentos de arquivo - dentro da área da ECM poderá ser bastante discutível, uma vez que estas são geralmente compreendidas como áreas relacionadas mas independentes e separadas, e que vão mais além do que a gestão da informação não-estruturada, pelo menos segundo a perspectiva da Ciência da Informação (Pinto e Silva, 2005). **No entanto, podemos também olhar para estas componentes como algo que não está necessariamente “dentro” da área de ECM, mas sim como algo que está diretamente relacionado com o âmbito da ECM e que deve ser envolvido nos processos do ciclo de gestão dos conteúdos (e que se mantém independentemente).** Esta última perspectiva parece-me mais realista e lógica.

- ✓ **Armazenar** – esta fase está relacionada com o armazenamento temporário da informação que ainda não está avaliada como sendo “para preservar” - apesar de as tecnologias utilizadas poderem ser as mesmas que as utilizadas que na fase “Preservar”, esta fase mantém-se separada precisamente por causa do carácter temporário do armazenamento da informação. Envolve os habituais repositórios, bases de dados e outros suportes de armazenamento. Importante referir a importância da meta-informação nesta fase, uma vez que será essencial para organizar a informação nos repositórios (Kampffmeyer, 2004; AIIM, 2012).
- ✓ **Preservar** – envolve o armazenamento da informação a longo termo, com utilização de backups e outros mecanismos de segurança e controlo de riscos. Este tipo de armazenamento contém informação que foi avaliada e considerada como sendo para arquivo e de preservação essencial. A informação preservada pode voltar mais tarde a ser avaliada, e ser eliminada ou transposta para um repositório temporário, caso a sua preservação não seja mais necessária (Kampffmeyer, 2004; AIIM, 2012).
- ✓ **Transmissão/distribuição (“deliver”)** – esta fase envolve os aspetos relacionados com a distribuição e apresentação da informação ao utilizador, ou seja com a gestão dos outputs (portais web, intranets, internet, etc.). Envolve processos de personalização de layouts, compressão da informação, segurança da informação disponibilizada, entre outros (Kampffmeyer, 2004; AIIM, 2012).

Vistas as várias fases, é **de notar que apesar da classificação ser referida na fase de captura, não é dada grande relevância há gestão de taxonomias, ontologias ou outro tipo de esquemas de classificação, o que se mostra negativo. A gestão e desenvolvimento das classificações é um dos processos mais importantes na adaptação dos sistemas ECM aos requisitos e contextos organizacionais, uma vez que tem um grande impacto na qualidade da organização dos conteúdos e, consequentemente, na recuperação dos mesmos.** Penso que seria importante refletir este aspeto na fase de “Gerir”, de forma a mostrar a importância e influência do mesmo nas outras fases do ciclo. Em alternativa, poderia ser incluída uma fase “Organizar”, onde seriam tidos em conta os aspetos relacionados com a gestão das classificações. Outro aspeto a referir é o facto das fases “Armazenar” e “Preservar” estarem separadas quando fundamentalmente são a mesma coisa, apenas se altera o período de tempo e o cuidado com que são armazenados os conteúdos. Ainda, seria importante valorizar a questão da avaliação dos conteúdos - de forma a perceber qual o repositório de armazenamento para cada conteúdo, e se estes seriam para armazenamento temporário, para preservação ou para eliminação -, algo que não é apenas referido brevemente na literatura que menciona este modelo de ciclo mais influente na indústria (Kampffmeyer, 2004; AIIM, 2012).

A proposta do modelo de ciclo da AIIM é bastante semelhante às propostas da Ciência da Informação para o ciclo de gestão da informação (ver figura 2⁵), o que é bastante positivo, uma vez que demonstra, de certa forma, a preocupação da indústria, das organizações e dos profissionais com o valor da informação e a necessidade de gestão da mesma desde a sua criação ou receção, até à sua transmissão ou preservação. Como podemos verificar na figura 2, todas as fases do ciclo de gestão de conteúdos estão representada, de uma ou de outra forma, no ciclo de gestão de informação proposto por Manuela Pinto e Armando Silva (2005), mesmo que em algumas delas seja necessário fazer algum esforço racional:

⁵ Ilustração 2 foi adaptada a partir do modelo de gestão da informação apresentado em: Pinto, M.A., Silva, A.M. (2005). “Um Modelo Sistémico E Integral De Gestão Da Informação Nas Organizações”. 2º Congresso Internacional de Gestão da Tecnologia e Sistemas de Informação, São Paulo, Brasil, acedido a 25 de outubro de 2012, <http://ler.letras.up.pt/uploads/ficheiros/3085.pdf>.

- ✓ **Capturar** – corresponde à fase 2 do ciclo de gestão da informação, no entanto segundo a descrição desta fase na proposta de ciclo analisada (Kampffmeyer, 2004; AIIM, 2012), também inclui muitos aspetos associados à fase 3, nomeadamente a indexação e classificação automática de informação;
- ✓ **Gerir** – corresponde fundamentalmente às fases de planeamento e de organização, no entanto pode envolver aspetos de todas as fases do ciclo de informação dada a quantidade de componentes a que está associada (segundo o ciclo de gestão de conteúdos que analisamos anteriormente);
- ✓ **Armazenar** – corresponde essencialmente à fase 5, no entanto, poderá envolver aspetos da fase 3 e 4 (tendo em conta que a organização e avaliação dos conteúdos é pouco referida no ciclo de gestão de conteúdos que analisamos);
- ✓ **Preservar** – corresponde à fase 7;
- ✓ **Transmitir** – corresponde diretamente à fase 6.

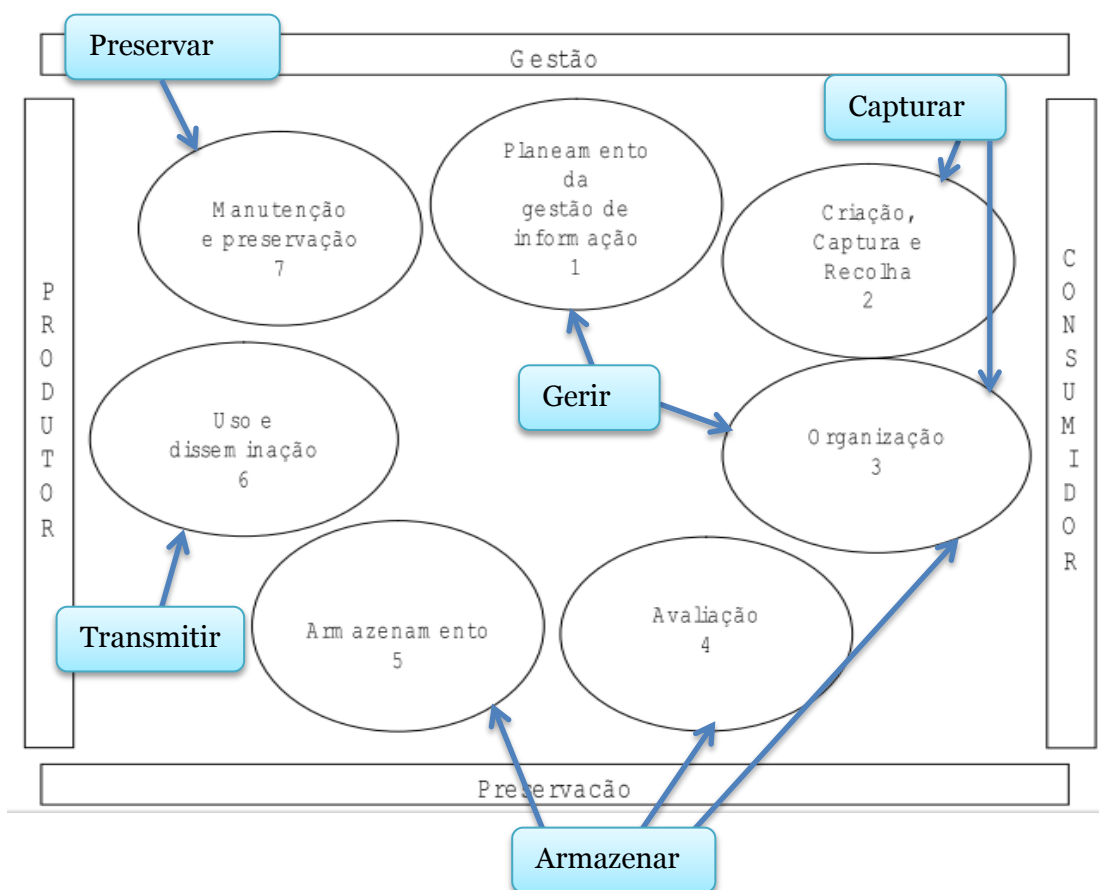


Ilustração 2 - Comparação das fases do ciclo de gestão de conteúdos (Kampffmeyer, 2005; AIIM, 2012) com o ciclo de gestão da informação (Pinto e Silva, 2005)

O desenvolvimento de classificação está sobretudo associado à fase “Gerir” do ciclo da gestão de conteúdos e, por comparação, à fase de planeamento e de organização da informação do ciclo de gestão da informação. Porém, as classificações têm importância em todas as fases dos dois ciclos comparados, assumindo especial influência durante as fases de captura, organização, armazenamento e uso/transmissão da informação, uma vez que estas são as fases onde os processos de classificação, organização e recuperação de informação estão mais vinculados.

Pode-se concluir que a ECM deve seguir uma perspetiva holística do ciclo de gestão de conteúdos, que integre estratégias, tecnologias, práticas e métodos para todas as fases de vida da informação, desde a sua criação ou recolha, até à preservação ou eliminação. Para privilegiar a boa organização e recuperação dos conteúdos, a ECM deve integrar processos de classificação que alcancem todas as fases do ciclo, para que a informação possa estar disponível rapidamente no suporte às atividades e processos organizacionais. Neste sentido, o desenvolvimento e implementação de classificações à medida dos requisitos e contextos organizacionais é essencial para a eficiência e eficácia deste ciclo.

Compreendido o âmbito e o posicionamento da classificação relativamente às várias fases e atividades dentro do ciclo de gestão dos conteúdos, é possível agora avançar para as questões mais conceituais e práticas sobre o conceito “classificação”, na tentativa de se definir uma visão teórica e prática que suporte e guie a construção do método.

1.1.3 Classificação

Desde que existe a possibilidade de registar conhecimento, o Homem sempre sentiu a necessidade de o organizar, de forma a o poder compreender e transmitir de

forma mais eficaz. Atualmente, os profissionais de informação funcionam como os intermediários entre o conhecimento registrado – a informação-, e os seus utilizadores, organizando e estruturando a informação de forma a poder ser recuperada de forma mais fácil e rápida. O conhecimento, por si só, não é estruturado - a sua organização e estruturação é imposta ao longo do tempo por filósofos e eruditos, de forma a facilitar a comunicação, o ensino e a aprendizagem, e a própria compreensão das várias áreas do conhecimento humano, ou de qualquer outra temática complexa o suficiente para “exigir” divisões do conhecimento (Iyer, 1995).

Tendo em conta a contínua evolução dinâmica do conhecimento, e a subjetividade inerente à sua estruturação (é imposta por humanos com visões e perspectivas diferentes), é possível concluir que não existe uma verdade absoluta que guie todas as estruturas do conhecimento, o que é evidente pelas visões divergentes entre áreas que estudam as mesmas realidades. Neste sentido, as divisões tradicionais do conhecimento eram estruturadas sempre com foco na especificidade de cada disciplina ou área, o que resultava no estabelecimento de fronteiras rigorosas entre as áreas. No entanto, a interdisciplinaridade e o dinamismo do conhecimento mostram que essa conceção tradicional era redutora e inadequada, uma vez que não demonstrava a complexidade do conhecimento humano, nem as relações existentes entre as várias áreas. É importante, então, ter em conta que o conhecimento e a informação estão integrados numa rede universal enorme, onde as ideias se encontram relacionadas e interligadas de várias formas, com o intuito de representar e interpretar o mundo de forma realista. Neste sentido, cada área do conhecimento está integrada numa rede e, conseqüentemente, associada a outras áreas do conhecimento – aliás, as áreas associadas a cada área do conhecimento são essenciais para a compreensão da própria natureza de cada uma destas, principalmente devido às relações dinâmicas que existem entre as mesmas (Iyer, 1995).

Posto isto, podemos dizer que a construção de uma estrutura para organização do conhecimento – como uma classificação-, não se deve cingir a uma visão redutora de uma única área ou disciplina do conhecimento – em vez disso, a finalidade, os objetivos e o âmbito da estrutura a ser criada devem guiar o processo do seu desenvolvimento, de forma a esta representar de forma realista o conhecimento importante e útil do contexto para o qual é direcionada. **Se transpusermos esta ideia para o mundo**

empresarial, podemos chegar à conclusão de que a estrutura do conhecimento de uma organização deve ser sobretudo baseada num âmbito determinado pela missão e contexto específico da empresa ou instituição, deixando para segundo plano as disciplinas científicas (que serão até mais representativas do conhecimento, mas que acabarão por não ser as mais realistas e ajustadas à natureza prática de um contexto empresarial). Assim, deve existir um compromisso entre o foco no contexto e características específicas do contexto para o qual a estrutura é concebida, e o suporte secundário dado pelas estruturas pré-definidas para a área – desta forma, a estrutura do conhecimento será representativa do contexto específico, mas ficará enquadrada e interligada numa rede de conhecimento.

Nesta secção do trabalho, iremos então tentar compreender as várias visões que existem relativamente à classificação, esclarecer a natureza do conceito dentro de um contexto mais prático e aplicado à gestão de conteúdos empresariais, e rever questões mais práticas como os princípios e os modelos estruturais de classificações. O resultado desta discussão será fulcral para a definição da visão teórica que guiará e suportará parte das decisões tomadas na construção do método, e dará vários apontadores para as opções mais práticas.

1.1.3.1 O conceito “Classificação”

Na perspetiva da Ciência da Informação, quando falamos de classificação, podemos estar a referir-nos a três conceitos distintos (Jacob, 2004):

- ✓ Um sistema de classes, construído e ordenado de acordo com um conjunto determinado de princípios e critérios definidos, e utilizado para organizar um conjunto de entidades;
- ✓ Uma classe ou grupo de classes num sistema de classificação (ex: “qual a classificação desta entidade?”);
- ✓ O processo de atribuição de entidades a classes de um sistema de classificação.

O conceito mais importante neste trabalho é, sobretudo, o primeiro, ou seja, o sistema de classes – é esta ferramenta que se pretende desenvolver à medida de requisitos e contextos organizacionais específicos, de forma a organizar eficazmente o conjunto de conteúdos da organização. No entanto, para se compreender como realmente funciona estes sistemas de classes é necessário compreender também a classificação como processo, uma vez que o sistema de classes influencia diretamente o processo de classificação, e vice-versa.

Este processo, tal como foi indicado acima, envolve a atribuição ordenada e sistemática de cada uma das entidades de um determinado conjunto a uma única classe pertencente a um sistema de classes. As classes deste sistema terão que ser, para isso, mutuamente exclusivas e não-sobrepostas, de forma a evitar que uma entidade pertença simultaneamente a duas ou mais classes. Desta forma, todas as entidades ou indivíduos pertencentes a uma classe compartilham o mesmo conjunto de características essenciais – esta similaridade dos membros de uma classe assegura que esta última pode ser definida pelo conjunto de características que os seus membros compartilham, e que a distingue de todas as outras classes da estrutura. **Este conjunto de características individualmente necessárias e conjuntamente suficientes é intrínseco a todas as entidades de uma classe e, por isso, define a essência da própria classe e do conceito que esta representa** (Jacob, 2004; Yier, 2005).

Este processo de classificação deve obedecer sempre ao mesmo conjunto de princípios que governou a construção, estruturação e ordenação do sistema de classes, para que o processo seja o mais coerente e sistemático possível.

1.1.3.2 Classificação, categorização e indexação

O facto do termo “classificação” ser um termo bastante utilizado fora do âmbito da Ciência da Informação, e de ser também considerado muitas vezes como sinónimo de indexação e categorização, resulta num uso indiscriminado dos três termos em muita da literatura científica e profissional. Cada um desses termos refere-se a um

conceitos e processos diferentes, e é importante, por isso, definirmos os processos de “categorização” e “indexação”.

A **categorização**, também como a classificação, é um processo que permite ordenar e organizar um grupo de entidades ou indivíduos através de agrupamento por similaridade, no entanto a forma como é feita esta ordenação é diferente do processo de classificação. Enquanto na classificação tradicional a atribuição de entidades é feita de forma rigorosa e exclusiva – ou seja, uma entidade é ou não é membro de uma determinada classe baseado num conjunto de princípios e critérios pré-determinados para aquela classificação -, no processo de categorização a atribuição é feita de forma flexível e são feitas associações no reconhecimento de quaisquer similaridades que existam entre as entidades, em qualquer momento. Desta forma, enquanto uma classificação organiza um conjunto de entidades num sistema de classes mutuamente exclusivas e não-sobrepostas de acordo com princípios pré-estabelecidos e fixos, a categorização organiza um conjunto de entidades em grupos nos quais os seus membros partilham uma qualquer similaridade dentro de um determinado contexto imediato. Este contexto pode variar e com eles também variam, consequentemente, os princípios para a organização e ordenação das entidades – e é este aspeto que dá flexibilidade ao processo de categorização. Neste sentido, na categorização não existe a questão das características necessárias e suficientes para uma entidade pertencer a uma classe, o que faz com que as fronteiras entre categorias sejam imprecisas (Jacob, 2004).

No que diz respeito ao processo de **indexação**, este pode ser encarado como sendo a representação de um documento ou conteúdo num índice com a finalidade de recuperação do mesmo. Uma definição alternativa, como a de Anderson e Pérez-Carballo 2001a) diz-nos que:

*“Indexar significa simplesmente apontar para ou indicar o conteúdo, significado, propósito ou características de mensagens, textos ou documentos”.*⁶

⁶ Anderson, J.; Pérez-Carballo, J. (2001a). ‘The nature of indexing: how humans and machines analyze messages and texts for retrieval. Part I: Research, and the nature of human indexing.’ *Information Processing and Management*, **37**(2), 233, acedido a 9 de janeiro de 2013, <http://comminfo.rutgers.edu/~muresan/IR/Docs/Articles/artAnderson2001a.pdf>.

A indexação é, portanto, também um processo essencial para a gestão de informação de uma organização, essencial para a recuperação de informação. O processo de indexação passa, então, pela análise e identificação do assunto de determinado objeto informacional e da sua representação através de um termo - este será o ponto de acesso que permitirá a recuperação do objeto informacional. Normalmente, os termos que são utilizados na representação dos assuntos e conceitos identificados nos objetos informacionais fazem parte de um vocabulário controlado, que não é mais do que uma lista organizada de termos autorizados e pré-selecionados pelo criador do vocabulário para serem utilizados como pontos de acesso dos objetos; no entanto, em alguns sistemas não existe qualquer tipo de restrição ao vocabulário a utilizar, e desse modo é utilizada a linguagem natural no processo de indexação. De referir que no âmbito da Ciência da Informação, subentende-se normalmente “indexação” como sendo indexação por assuntos apenas, no entanto vários autores, tal como Anderson e Pérez-Carballo, falam também da descrição de características de documentos e informação (tais como autor, descritor, características físicas, etc.).

São utilizadas duas abordagens para o processo de indexação: por examinação humana ou por algoritmo computacional, ou, por outras palavras, indexação humana ou indexação automática, respetivamente. A indexação automática surgiu muita por culpa da explosão informacional crescente que foi pedindo (e continua a pedir) métodos mais rápidos e eficazes para indexar a quantidade de informação que continua a crescer exponencialmente. A indexação automática é atualmente essencial para uma organização conseguir atribuir meta-informação à quantidade de informação que cria e recolhe diariamente, devido sobretudo à sua rapidez de processamento. Além desta vantagem, os métodos automáticos de indexação são ainda mais previsíveis e consistentes (o que contribui para a objetividade que se pretende nesta tarefa), mais baratos e eficientes, têm menor probabilidade de erros, auxiliam o utilizador através do processamento de associações entre os objetos informacionais, entre outras mais-valias; por outro lado, são pouco flexíveis (executam sempre as tarefas da mesma forma), podem ter problemas de precisão na análise de documentos com características singulares, necessitam de atualizações frequentes e, pois claro, precisam de pessoal com know-how tecnológico para serem utilizadas da melhor forma. Apesar das inúmeras vantagens, não podemos dizer indiscutivelmente que a abordagem automática à indexação é melhor que a abordagem humana; acima de tudo, as duas

abordagens complementam-se e devem ser utilizadas em conjunto para aproveitar os pontos fortes de cada uma.

Vistas as características que definem o processo de indexação, e dado o uso indiscriminado dos termos “indexação” e “classificação”, põe-se novamente a questão de qual será efetivamente a diferença entre estes dois processos. Em parte, a indexação é um conceito muito menos restrito e rigoroso que a classificação, uma vez que designa apenas a associação de uma série de palavras-chave (os pontos de acesso) e descritores a um objeto informacional, elementos que farão parte da meta-informação deste; pelo contrário, o processo de classificação, tal como vimos acima, envolve a organização e agrupamento de entidades (que neste caso seriam objetos informacionais) em classes mutuamente exclusivas - neste caso, a classe seria também parte essencial da meta-informação que permitirá recuperar o objeto informacional. Porém, Wellisch (citado em Hjørland, 2010b) faz um contra-argumento bastante interessante a esta ideia, mencionando que se for utilizado na indexação um tesauro baseado numa classificação facetada, a diferença entre o processo de classificação e indexação não é notória. Além disso, Wellisch afirma ainda que a principal diferença na organização da informação não é entre o processo de indexação e classificação, mas sim entre a utilização de sistemas baseados em vocabulários controlados e baseados em vocabulários não-controlados – e segundo este argumento, o autor conclui que, no fundo, o ato de classificação é também um ato de indexação e vice-versa. No entanto, podem ser apontadas duas lacunas a este argumento de Wellisch, nomeadamente:

- ✓ O facto das classes dos sistemas de classificação tradicionais serem mutuamente exclusivas e não-sobrepostas não “encaixa” com o processo de indexação, uma vez que neste último podem ser associadas várias palavras-chave ao mesmo objeto informacional, o que faz com que o objeto pertença a múltiplas “classes”;
- ✓ A indexação automática é baseada em técnicas de análise de texto que nem sempre indicam objetivamente o principal assunto do objeto informacional, que seria essencial para determinar a classe deste último.

Apesar destas duas lacunas, o argumento de Wellisch não é irrealista de todo. É possível imaginar num processo de indexação complementado por uma classificação tradicional, onde os mecanismos automáticos indexam o objeto informacional apenas

com uma única palavra-chave representativa de uma classe – nesta visão, os processos de indexação e classificação seriam, na prática, bastante semelhantes. **Pode-se dizer então que, na prática, o processo de classificação é também um processo de indexação: apesar do processo de classificação ser mais rigoroso, as classes de um sistema de classificação, para além de serem utilizadas para organizar um domínio, também são utilizadas para representar o conteúdo de um objeto informacional para fins de recuperação de informação. Portanto, a distinção não é causada por supostas diferenças fundamentais e práticas entre os dois processos, mas sim pela utilização de vocabulários controlados ou de classificações na fase de representação do conteúdo dos objetos informacionais.** Assim, podemos concluir que o processo de indexação é muito importante para a classificação e deverá ser tida em conta durante o desenvolvimento e gestão de classificações.

Compreendida a conceção tradicional do processo de classificação, e a sua relação com conceitos relacionados como a indexação e a categorização, podermos agora passar à descrição dos principais princípios que estão na base de um sistema de classes.

1.1.3.3 Princípios da classificação

Os princípios e critérios que estão na base de cada classificação (agora como sistema, e não processo) são arbitrários, isto é, são definidos de acordo com e refletem uma única visão do domínio das entidades. Este facto faz com que o desenvolvimento de um sistema de classes seja um processo complexo, que varia consoante o âmbito e a finalidade que se pretende dar ao produto final. No entanto, a conceção mais tradicional de classificação fornece uma série de princípios normativos que se mostram úteis para a descrição deste processo. Estes princípios pretendem também ajudar a que se possa manter a estabilidade da estrutura das classificações ao longo de tempo, e permitindo simultaneamente que as mesmas possuam um carácter flexível, eficiente e intuitivo.

Iyer (2005) encontra três níveis universais no processo tradicional de classificação:

- ✓ **Nível concetual**, onde se selecionam os conceitos ou entidades, e a forma como estes serão estruturados;
- ✓ **Nível terminológico**, onde são definidos os termos representativos dos conceitos;
- ✓ **Nível da notação**, onde é atribuído um código único a cada entidade da estrutura.

Para cada um destes níveis, Iyer define uma série de princípios que devem conduzir as tarefas enquadradas em cada um dos mesmos.

No **nível concetual**, o processo de desenvolvimento de uma classificação envolve a divisão do conhecimento em grupos, que por sua vez serão divididos em partições menores, até que se atinja o nível de especificidade desejado. A **seleção das características para a divisão das classes**, e a ordem pelo qual estes são aplicados, são dois dos aspetos essenciais do processo de desenvolvimento de uma classificação, definindo a sua estrutura. No processo de seleção das características para a divisão, existem alguns princípios que devem ser tidos em conta, e que tornam as divisões de um sistema de classes mais rigorosas que aquelas que utilizamos para categorizar e organizar mentalmente o conhecimento e a realidade:

- ✓ As divisões entre as classes devem ser claras, refletindo características e diferenças evidentes;
- ✓ As divisões entre as classes devem ser permanentes, tanto quanto possível. Nesse sentido, as características selecionadas para dividir as classes devem ser estáveis ao longo do tempo (por exemplo, formar grupos de países tendo em conta a população não será uma divisão estável, enquanto dividir por continente já o será);
- ✓ As características selecionadas para a divisão de classes devem ser verificáveis por todos os utilizadores futuros. Deve-se sempre optar pelas características mais acessíveis, no entanto sem simplificar demasiado a classificação de forma a não prejudicar a sua utilidade futura;

- ✓ As características selecionadas devem ser relevantes para a finalidade da classificação.

A complexidade na organização do conhecimento exige que sejam definidas várias características, de forma a formar subclasses dentro de cada classe, e assim sucessivamente. No entanto, em cada nível da estrutura, apenas uma característica deve ser utilizada, de forma a manter as classes mutuamente exclusivas.

A **sequência das características** que determinam as divisões, ou seja, a ordem pelo qual as características são utilizadas para agrupar e separar os conceitos, é também importante, e deve ser escolhida de modo a ir de encontro às necessidades do maior número de utilizadores. Por exemplo, se a maioria dos utilizadores de uma loja de música preferir que os álbuns sejam preferencialmente divididos por “género musical”, a classificação deverá primeiro aplicar essa característica na divisão e só depois dividir por “nacionalidade”. Esta ordenação é sobretudo importante em classificações organizadas apenas numa única estrutura hierárquica - nas classificações facetadas este problema é reduzido, como poderemos ver mais à frente neste trabalho. Importante é também ter em conta que a sequência das características a serem aplicadas deve ser mantida em toda a classificação, de modo a manter a estrutura hierárquica consistente.

As divisões fazem com que cada nível do sistema de classes contenha uma série de **classes coordenadas**, ou seja, equivalentes no que diz respeito ao nível na estrutura e pertencentes à mesma classe subordinante. Existem dois aspetos essenciais a ter em conta no que diz respeito à organização das classes coordenadas:

- ✓ As classes coordenadas devem acomodar todos os conceitos possíveis nesse nível da estrutura. Por exemplo, dentro de uma hipotética classe “estações do ano”, “outono”, “inverno”, “primavera” e “verão” acomodariam todos os conceitos possíveis. No entanto, devido à natureza evolutiva e dinâmica do conhecimento, muitas vezes é impossível saber e definir todos os conceitos pertencentes a uma classe – nestes casos, a identificação de uma classe do tipo “outros” ou “miscelânea” é útil para cumprir este requisito. Caso um novo conceito seja encontrado e não se

adeque a nenhuma das classes definidas, deverá ser “acomodado” nesta classe de “outros”, até que se crie uma classe própria para o mesmo;

- ✓ A ordenação das classes coordenadas em cada nível também é importante para a organização dos conceitos. Esta ordenação, quando possível, deve ser feita de acordo com um princípio pré-estabelecido, como por exemplo:
 - Tempo ou sequência evolutiva (por exemplo, para organizar conceitos que são definidos por um próprio período de tempo);
 - Sequências posicionais (por exemplo, para ordenar os planetas numa classificação);
 - Fases de um processo (por exemplo, para organizar os conceitos relacionados com o processo de construção de um carro);
 - Grau de complexidade;
 - Dimensão;
 - Ordem alfabética.

No **nível terminológico**, é importante que a escolha dos termos para representar os conceitos seja feita de acordo com as características do contexto em que será utilizada a classificação, particularmente o tipo de utilizadores ao qual esta é direcionada. Além deste aspeto, existem outros princípios gerais, apontados por Iyer:

- ✓ Deve ser sempre utilizado o termo mais atual, correto e neutral para representa o conceito. Apesar disso, devem ser tidas sempre em conta características do contexto para o qual é direcionada a classificação, uma vez que poderão existir determinados termos específicos e “invulgares” que sejam mais utilizados dentro do contexto particular;
- ✓ Deverão ser utilizadas as formas mais simples dos termos, uma vez que a própria estrutura hierárquica providencia o contexto para cada termo.

Quanto ao **nível da notação**, podemos dizer que é sobretudo importante quando existe a necessidade de organizar fisicamente os objetos. No entanto, a notação - o sistema de códigos associados a cada conceito da classificação - poderá ser também útil no momento da pesquisa eletrónica da informação. Na conceção tradicional da classificação, existem vários princípios e regras para a elaboração de notações, definidos sobretudo para normalizar a organização dos documentos físicos em bibliotecas e arquivos, entre os quais (Iyer, 2005):

- ✓ A notação de cada classe deve ser única;
- ✓ As notações devem obedecer e refletir a estrutura hierárquica de classes;
- ✓ O sistema de notações deve ser elaborado de forma a permitir atualizações futuras;
- ✓ A escolha de novos conceitos para a classificação deve governar a atribuição da notação, e não o contrário (mesmo que isso exija a revisão das notações);
- ✓ As notações atribuídas devem ter em conta a familiaridade dos utilizadores com o código utilizado nas mesmas;
- ✓ As notações devem ser criadas tendo em conta um balanço da sua brevidade, individualidade e eficiência;
- ✓ Quando possível, as notações devem ser elaboradas de forma a refletir o conceito ao qual é atribuída, de forma a torná-la mais intuitiva e a promover a memorização por parte dos utilizadores.

No entanto, em ambiente eletrónico, a questão da organização física dos documentos não se coloca. Além disso, a natureza da pesquisa nos motores de busca é baseada na meta-informação atribuída a cada objeto informacional, e não num posicionamento físico normalizado. Assim, e tendo em conta que a gestão de conteúdos é sobretudo feita em portais eletrónicos, podemos afirmar que a organização, pesquisa e recuperação eletrónica de conteúdo não exige essencialmente a associação de uma notação a cada termo, no entanto a utilização da mesma revela-se útil em certos casos (como, por exemplo, para a ordenação das classes na estrutura).

É necessário ter em conta que os princípios apresentados acima estão integrados na perspetiva tradicional da classificação, centrada na organização do conhecimento em bibliotecas e arquivos, numa perspetiva de representação do conhecimento universal. Como tal, é essencial referir que, numa perspetiva de aplicação das classificações aos processos organizacionais, estes princípios funcionam como um auxílio aos profissionais, e não como regras rigorosas e inflexíveis. Assim, estes princípios da conceção tradicional devem ser compreendidos como base inicial no suporte à construção de classificações, e não como dogmas na aplicação de classificações a contextos organizacionais específicos.

Tal como foi dito, os princípios das classificações ajudam a manter a estabilidade e funcionalidade das classificações, permitindo com que seja possível desenvolver de forma coerente vários tipos de estruturas de classificação.

1.1.3.4 Modelos estruturais de sistemas de classificação

A categorização e a classificação são fundamentalmente processos mentais, através dos quais é possível organizar a informação no cérebro e estruturar a memória através de associações e grupos de categorias. É, pois, natural que as estruturas típicas das classificações reproduzam o processo mental lógico que o ser humano realiza para assimilar conhecimento e organizá-lo na sua memória – não só torna mais fácil o processo de criação da própria estrutura, mas também faz com a sua compreensão e interpretação pelos utilizadores seja lógica e intuitiva. As estruturas - ou esquemas – de classificação devem ser capazes de organizar os vários conceitos que se pretende representar, sempre numa perspectiva que privilegie a compreensão lógica e a utilidade e usabilidade da estrutura no contexto específico para o qual foi criada. O facto de cada tipo de classificação ter finalidades diferente resulta na existência de várias abordagens ao processo de estruturação do sistema de classes - o que por sua vez origina classificações com propriedades estruturais distintas, que dão a cada tipo de classificação diferentes pontos fortes e fracos. De seguida, são descritas as estruturas mais utilizadas na organização da informação, não aquelas cuja essência está mais ligada à conceção mais tradicional da CI, mas também outras cuja origem está ligada à influência de áreas como a Inteligência Artificial, e ao desenvolvimento das redes sociais e dos comportamentos informacionais associados à Internet.

Os modelos estruturais de classificação mais tradicionais são (Kwasnik, 1992; 1999):

- ✓ **Hierarquia** – as hierarquias são as estruturas preferidas para representar e organizar o conhecimento. A sua organização baseia-se na divisão de um “todo” nas suas partes – isto é, classes e subclasses -, seguindo sempre um sistema ordenado e sistemático de regras de associação e diferenciação entre conceitos a representar. Os requisitos estruturais das hierarquias incluem:

- Relações do tipo “é um” – numa hierarquia, as subclasses são sempre exemplos das classes diretamente subordinantes, e assim sucessivamente, à semelhança do que acontece nas classificações biológicas (como exemplo, e de forma muito genérica, “gato” é um tipo de “felino”, e “felino” é um tipo de “mamífero”);
- Herança – assegura que tudo o que é verdadeiro para as entidades de uma classe, também é verdadeiro para as entidades das suas classes subordinadas;
- Transitividade – como as características são herdadas pelas classes subordinadas, os membros das classes subordinadas são membros de todas as classes diretamente subordinantes (Se “gato” é um tipo de “felino”, e “felino” um tipo de “animal”, então “gato” é também um tipo de “animal”);
- Especificidade e profundidade - as classes subordinadas devem ter sempre mais especificidade e profundidade que as classes diretamente superiores na estrutura hierárquica. Utilizando os termos de Iyer (2005), as classes subordinadas deverão ser menos extensas e mais intensas que as respetivas classes subordinantes – ou seja, devem representar menos exaustivas que a classe diretamente superior na hierarquia, mas mais específicas.

Segundo Kwasnik (1999), as hierarquias têm vários pontos fortes que as tornam excelentes para representar domínios onde a natureza dos conceitos está bem definida e apresenta fronteiras claras. No entanto, o facto de ver o domínio apenas segundo um ponto de vista (relações do tipo “é um”) torna-a insuficiente para representar as diferentes perspetivas que poderemos ter do mesmo conceito.

- ✓ **Árvore** – a estrutura em árvore é em grande parte semelhante, no entanto não utiliza a genérica relação “é um” que define a essência das hierarquias. Por causa deste aspeto, também não assumem as regras de herança de características. Assim, pode-se dizer que as classes estão organizadas por relações sistemáticas, mas não a genérica “é um” (por exemplo, dividir um “automóvel” pelas partes constituintes – “motor” e “carroçaria” não são tipos

de “automóvel”; poe-se dizer que se utiliza aqui uma relação do tipo “é parte de”). Este tipo de estrutura tem como principal ponto forte o facto de permitir que seja dada relevância a determinado tipo de relação entre as classes, permitindo outro ponto de vista do domínio que salienta determinada característica das entidades (Kwasnik, 1992). No entanto, e tal como as hierarquias, apenas permite um ponto de vista do domínio, e, além disso, a estrutura em árvore faz com que as classes coordenadas (no mesmo nível e dentro da mesma classes subordinante) possam não ter qualquer relação aparente (“motor” não tem relação aparente com “carroçaria”). Deste modo, este tipo de estrutura foca-se apenas nas relações verticais.

As **classificações facetadas**, não são, por natureza, diferentes modelos estruturais, mas sim uma abordagem diferente à classificação. Os trabalhos de Vickery e Ranganathan (citados em Tennis, 2011 e Broughton, 2011) foram bastante influentes no desenvolvimento desta abordagem. Ranganathan, na década de 1930, (citado em Broughton, 2006) é considerado geralmente como o criador do conceito de análise por facetas sua *Colon Classification*, onde propôs uma série de facetas fundamentais que seriam suficientes de representar qualquer fenómeno. Mais tarde, nos anos 50, o Classification Research Group foi formado e liderou a investigação neste tipo de abordagem, particularmente através de um dos seus membros, Brian Vickery (citado em Gnoli, 2011, Tennis, 2011, e Broughton, 2011). As classificações facetadas permitem representar vários pontos de vista do mesmo conceito, partindo assim do pressuposto que qualquer domínio é dinâmico e pode ser visto de diferentes formas e segundo diferentes perspetivas (por exemplo, um objeto pode ser visto segundo a sua forma, cor, função, etc.,). Cada faceta funciona como uma categoria, ou seja como uma classe numa hierarquia ou árvore (Iyer, 2005) – cada objeto é classificado em cada uma das facetas, de acordo com as características que cada faceta privilegia. As facetas resolvem assim um dos principais pontos fracos das hierarquias e das árvores, nomeadamente o inconveniente de estas conseguirem apenas representar um único ponto de vista do domínio. Esta abordagem implica, deste modo, a escolha de facetas e a análise das entidades sob o ponto de vista de cada uma destas. A informação em ambiente digital também se adequa à capacidade descritiva que a classificação facetada providencia – em ambiente digital a necessidade não está em organizar o material “fisicamente”, mas sim em descrevê-lo adequadamente de forma a suportar a recuperação, a navegação e a

apresentação de resultados (Broughton, 2006). Este tipo de modelo estrutural apresenta, então, várias vantagens (Kwasnik, 1999; Broughton, 2006):

- ✓ Flexibilidade – o modelo por facetas permite que cada objeto seja descrito através de vários atributos independentes, o que irá permitir que o utilizador possa combinar atributos e recuperar os objetos que se ajustem a essa combinação. Esta abordagem chama-se pós-coordenação, e significa que a combinação dos vários atributos é feita no momento da recuperação, ao contrário das pré-coordenadas, onde o próprio indexador tem de combinar os vários atributos. O utilizador pode assim relacionar livremente os termos das várias facetas. Esta abordagem torna a classificação por facetas muito mais dinâmica e multidimensional;
- ✓ Expressividade – as facetas permitem que os objetos sejam definidos na classificação com maior detalhe e precisão, e também de forma mais sintética;
- ✓ Diversidade de utilização – o modelo estrutural por facetas ajusta-se a outras técnicas de pesquisa, tal como a navegação ou browsing ⁷(Hjørland, 2011;). A visualização das facetas através de menus (como se constata em vários sites) também permite que o utilizador perceba de forma rápida de que forma os conteúdos estão categorizados e organizados. Este tipo de navegação, aliado à estrutura por facetas, ajusta-se perfeitamente aos comportamentos de pesquisa na Internet, permitindo que o utilizador escolha a sua forma de pesquisa;
- ✓ Estrutura dinâmica – este tipo de classificação não exige tanto rigor teórico no momento da criação, uma vez que não possui uma estrutura pré-definida;
- ✓ Capacidade de atualização – como a classificação por facetas apenas exige a definição de facetas, a construção da classificação pode ser feita ad hoc, desde que se mantenham as facetas se mantenham estáveis. A cada novo objeto, podem surgir novos atributos para cada uma das facetas, o que torna relativamente fácil a atualização da classificação;

⁷ *Browsing* refere-se ao processo de visualizar páginas, uma de cada vez, navegando entre eles de forma sequencial através de *hyperlinks*. É uma técnica de pesquisa onde não há a procura de um resultado específico, mas sim a procura da informação disponível (ao contrário da pesquisa normal através da introdução de uma interrogação de pesquisa ou *query*) (Park, 2011)

- ✓ Compatibilidade – este tipo de estrutura pode ser facilmente convertido num tesouro, ou nouro tipo de vocabulário controlado ou índice.;
- ✓ Perspetivas múltiplas – cada objeto pode ser visto sob várias perspetivas e definido de formas distintas, o que é uma se traduz numa visão mais fiel das realidades;

Apesar do dinamismo, multidimensionalidade e flexibilidade desta abordagem, Kwasnik (1999) encontra alguns pontos fracos:

- ✓ Dificuldade em estabelecer facetas adequadas ao domínio – a qualidade deste tipo de classificação está dependente da adequação das facetas ao domínio e às necessidades dos utilizadores. Assim, o estabelecimento de facetas está dependente do conhecimento do domínio e dos utilizadores, o que pode ser difícil de compreender;
- ✓ Ordem de citação – se a classificação utilizar algum tipo de notação, poderá tornar-se demasiado complexo enunciar e ordenar as notações de cada faceta.

Existem alguns sistemas de “facetas fundamentais”, propostos por Ranganathan e Vickery, que se afirmam como aplicáveis a qualquer domínio, e que são capazes de ajudar a diminuir a dificuldade no estabelecimento de facetas. No entanto, esses sistemas podem-se revelar demasiado técnicos, gerais e científicos para uma aplicação num contexto empresarial específico. Desta forma, as facetas que ambos propõem para a representação da realidade devem ser vistas como um apoio ao desenvolvimento de classificações para contextos específicos, e não como uma regra, sob pena de se tornar o processo de construção da classificação demasiado complexo, e de prejudicar a sua usabilidade.

Os modelos estruturais descritos nas últimas páginas, baseados na conceção mais tradicional da classificação, podem ser divididos em dois grupos: esquemas enumerativos ou não-facetados (hierarquia e árvore) e esquemas facetados (classificação facetada). No entanto, a influência da evolução tecnológica – particularmente a área da inteligência artificial –, e os novos comportamentos informacionais provocados pela explosão da internet, têm contribuído para o

surgimento de novas estruturas para organizar o conhecimento. Destas, as ontologias e as folksonomias são as mais utilizadas:

- ✓ **Folksonomia** – é um tipo de “classificação web”, termo utilizado para descrever sistemas de informação na web que incorporam categorias (Park, 2011). A estrutura das folksonomias baseia-se em “tags”, um tipo de palavras-chave que os utilizadores da Web podem atribuir a objetos na Internet, com o intuito de os partilhar, organizar e recuperar mais tarde – o processo de atribuição de “tags” chama-se “tagging” e a folksonomia emerge deste. Esta consiste, assim, num espaço virtual plano definido pelo conjunto de tags utilizadas por um grupo de utilizadores para categorizar recursos web, e é normalmente representada através de uma “nuvem” (ou *cloud* (fig.3⁸)) de tags, onde é dado destaque às tags mais utilizadas. Devido à liberdade que é dada aos utilizadores no momento da atribuição de tags, este tipo de estrutura enquadra-se mais no conceito de “categorização” do que de “classificação” tradicional, uma vez que as classes não serão necessariamente mutuamente exclusivas e não sobrepostas - o utilizador categoriza os recursos tendo como base o que bem entenderem, o que faz com que cada objeto possa pertencer a várias categorias. Contudo, têm sido propostas algumas técnicas para aproximar os vocabulários controlados das tags criadas por utilizadores, de forma a melhorar a eficácia de pesquisa. Devido às características deste tipo de estrutura, ela é mais direcionada para a técnica de pesquisa por browsing, e não por pesquisa direta. Este tipo de estrutura tem as seguintes vantagens (Park, 2011; Mai, 2011a)):



Ilustração 3 - Tag cloud

⁸ Ilustração 3 retirada de Vikiinfos Blog, acedido a 30 de março, <http://vikiinfos.blogspot.pt/2011/10/how-to-add-tag-cloud-to-your-blog.html#axzz2OyqakvsX>

- Colaboração e partilha de conhecimento – os utilizadores são encorajados a participar publicamente na partilha de informação, motivando a partilha de conhecimento e a contribuição na organização da informação. É a principal vantagem das folksonomias, permitindo que se criem verdadeiras comunidades;
- Usabilidade – qualquer utilizador pode atribuir tags aos objetos, não apenas os autores ou especialistas do domínio (não existe uma posição autoritária sobre o sistema). Isto faz com que a folksonomia seja útil para qualquer pessoa, uma vez que é possível ser usada numa perspetiva individual;
- Interação – os utilizadores podem consultar as tags atribuídas pelos outros utilizadores a determinado recurso e optar por alterar a sua própria tag que havia atribuído ao mesmo recurso - criando-se assim, uma categoria comum. A interação entre as categorias idiossincráticas e as categorias comuns resultam numa visão pessoal e de comunidade na folksonomia;
- Pluralidade de perspetivas – o facto de cada pessoa poder contribuir com tags permite que cada um possa dar a sua perspetiva pessoal do recurso através da atribuição de tags à sua escolha. Isto faz com que o sistema não se torne enviesado por determinada perspetiva autoritária, representando fielmente as características, os interesses e o conhecimento de uma comunidade;
- Navegabilidade – o facto de cada “tag” ser não só uma categoria, mas também um link para objetos com essa categoria associada, permite que os utilizadores possam fazer browsing de forma intuitiva e descobrir informação relevante para as suas necessidades;
- Capacidade de atualização - novas tags podem ser adicionadas a qualquer momento sem qualquer tipo de problema.

As principais desvantagens são:

- Falta de estruturação – as folksonomias não apresentam uma estrutura visível, como uma hierarquia, sendo basicamente um conjunto de tags livre na “cloud”;
- Ambiguidade – a liberdade que é dada aos utilizadores resulta na existência de categorias com vários significados e de várias categorias diferentes com o mesmo significado, prejudicando a recuperação;
- Liberdade total – mais uma vez, a liberdade que é dada aos utilizadores poderá resultar em vários tipos de erros, como tags ortograficamente incorretas, ou uso descuidado de tags, o que se poderá revelar um caos caso não haja algum tipo de moderação.

As folksonomias permitem então que todos os utilizadores possam participar na organização do conhecimento, resultando numa grande quantidade de perspetivas e interpretações acerca de cada objeto informacional, o que se traduz numa evolução do conhecimento que temos acerca da informação (Mai, 2011a). Este tipo de estrutura baseada numa abordagem social e construtivista representa, assim, uma alternativa viável às tradicionais estruturas construídas numa base autoritária – aliás, o interessante será perceber como estas duas abordagens poderão ser utilizadas em conjunto de forma a aproveitar as mais-valias de ambas.

- ✓ **Rede semântica** – as redes semânticas, apesar de serem já utilizadas pela filosofia, psicologia e linguística, começaram a ter uma maior impacto com a sua implementação em sistemas tecnológicos pela área da Inteligência Artificial (Sowa, 2006), de forma a facilitar a partilha e a reutilização de conhecimento. Esta aplicação traduz-me normalmente na criação de ontologias, no entanto existem estruturas mais básicas e informais baseadas em redes semânticas, como por exemplo os mapas de conceitos. No entanto, as ontologias são o melhor exemplo para explicar as características das redes semânticas. Baseado na literatura sobre ontologias, podemos dizer que as redes semânticas são estruturais formais gráficas representando uma especificação explícita de uma

concretização de um domínio, providenciando uma compreensão comum deste último que pode ser comunicada entre utilizadores e as aplicações de sistemas (López, 2005). A principal característica destas estruturas é o facto de elas entrelaçarem a semântica formal das aplicações tecnológicas com a semântica natural compreensível pelos humanos – por exemplo, uma ontologia normal inclui objetos individuais ou instâncias, conceitos ou classes do domínio, as relações entre estes, as propriedades cruciais de cada conceito através de atributos, uma série de restrições e regras para restringer a interpretação, e axiomas que permitem fazer asserções sobre o domínio através da utilização de lógica formal (Gruber, 1993). São então redes semânticas que se traduzem em bases do conhecimento, suportando o funcionamento de sistemas “inteligentes”. As redes semânticas estão assim interação entre aplicações e utilizadores através da lógica formal, conseguindo descrever um domínio de forma mais complexa devido aos vários componentes que utiliza para fazer permitir a realização de asserções acerca de um domínio. A complexidade das ontologias faz com que a sua utilização seja direcionada para cenários mais complexos que os da gestão de conteúdos, tais como atividades de modelação de domínios e de comunicação entre sistemas e humanos através de inteligência artificial. Desse modo, o desenvolvimento de uma ontologia para um sistema de gestão de conteúdos não seria o mais eficiente. No entanto, os mapas de conceitos

(fig.4⁹) (Novak e Cañas, 2006) são estruturas mais simples que apenas utilizam conceitos e relações, serão úteis para sintetizar e esquematizar o

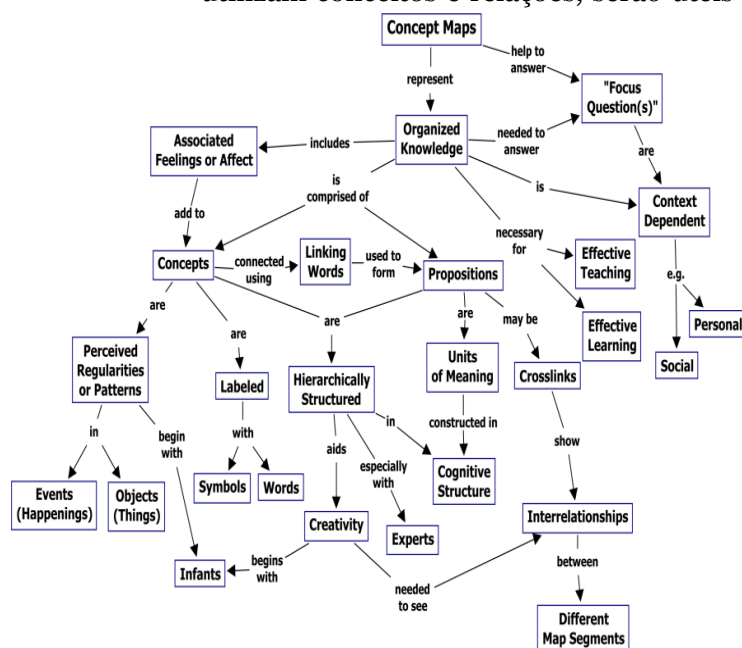


Ilustração 4 - Exemplo de mapa de conceitos

conhecimento de um domínio empresarial de forma eficiente e eficaz, podendo ser utilizados como base para classificações de conteúdos.

As últimas páginas apresentaram alguns dos principais modelos estruturais, cada um com a sua própria perspectiva da

organização do conhecimento. As

ontologias, especialmente, permitiram compreender a influência das áreas mais tecnológicas no conceito de “classificação”, através da introdução de uma nova perspectiva de interação entre sistema e utilizador humano. Neste sentido, na próxima secção é discutida a relação entre a classificação e a Inteligência Artificial.

1.1.3.5 Classificação para além da visão tradicional da Ciência da Informação

Nas últimas páginas foram analisados os aspetos principais da conceção tradicional de classificação, influenciada principalmente pela área da Ciência da Informação que está mais ligada à gestão da informação em bibliotecas e arquivos. No entanto, a Ciência da Informação, como área transdisciplinar, inclui no seu âmbito de estudo vários aspetos que não exclusivos da sua área, sendo também estudadas noutras

⁹ Ilustração 4 retirada de Novak, Joseph e Cañas, Alberto (2008). “The Theory Underlying Concept Maps and How to Construct and Use Them”. Technical Report IHMC Cmap Tools, Florida Institute for Human and Machine Cognition, acedido a 25 de março de 2013, <http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf>.

áreas, como a engenharia e a inovação tecnológica. A classificação é um desses aspectos com relevância noutras áreas para além da CI, nomeadamente na Inteligência Artificial.

A Inteligência Artificial (IA) emergiu como área de estudo no final dos anos 60, principalmente por influência do trabalho de John McCarthy e seus colegas, no entanto o conceito de “máquinas inteligentes” já era mencionado na mitologia grega (Buchanan, 2005). A IA está bastante ligada aos princípios da classificação e da representação do conhecimento, princípios que são utilizados como base de sistemas tecnológicos “inteligentes” (Iyer, 1995). Estes sistemas são, então, suportados por bases de conhecimento complexas que se vão desenvolvendo ao longo do tempo, muito à semelhança das classificações tradicionais utilizadas na área da Ciência da Informação para gestão da informação em bibliotecas. Estas bases de conhecimento, conjuntamente com uma série de algoritmos, permitem que estes sistemas possuam “inteligência” para raciocinar, fazer inferências e tomar decisões, de uma forma relativamente parecida ao que o cérebro humano é capaz de fazer – desta forma os sistemas não só ajudam a complementar o poder do cérebro humano, mas também permitem que saibamos mais sobre a forma como o ser humano raciocina e toma decisões.

A capacidade de raciocínio um sistema “inteligente” exige bases de conhecimento mais complexas que as típicas estruturas tradicionais da Ciência. Tal como foi possível perceber na última secção durante a análise das redes semânticas, as bases de conhecimento são compostas por mais elementos que apenas simples classes com relações básicas entre elas – pelo contrário, as bases utilizados requerem componentes como regras, relações com vários significados, instâncias de conceitos, propriedades descritoras, axiomas e restrições. Estas bases constituem redes semânticas que simulam o processo mental humano. Esta complexidade possibilita que o sistema possua capacidade de raciocínio, que lhe permite ir além das funções de representação do conhecimento proporcionadas pelas classificações tradicionais de CI. Entre estas funções e capacidades mais complexas (além da representação do conhecimento), encontramos as seguintes (McCarthy, 2007):

- ✓ Tomada de decisões – tomada de decisões baseada na linguagem matemática lógica, onde o programa escolhe a ação mais apropriada a determinado objetivo;

- ✓ Reconhecimento de padrões – comparação de observações com padrões definidos;
- ✓ Descoberta de relações – as regras de associação são capazes de descobrir correlações interessantes entre conceitos nas redes semânticas (Kamruzzaman, 2010);
- ✓ Inferência – capacidade de inferir novos factos a partir de outros, utilizando não só métodos indutivos, mas também outros mais complexos;
- ✓ Aprendizagem por experiência – os sistemas podem aprender através de regras e leis expressas na lógica da rede semântica;
- ✓ Planeamento – capacidade parecida com a tomada de decisões, onde o sistema é capaz de gerar estratégias para alcançar objetivos, baseando-se nos factos expressos através da semântica;
- ✓ Previsão – análise preditiva da classe em que se enquadra cada nova observação, baseando-se nos atributos desta, ou análise preditiva de um atributo desconhecido de uma observação (Ai Access, 2013). Além destas funcionalidades enquadradas de classificação, os sistemas “inteligentes” podem ainda ser capazes de fazer previsões sobre o futuro, através de análise estatística, prospeção de dados (*data mining*), e outros tipos de exploração de grande quantidade de dados. Esta funcionalidade envolve modelos de probabilidade, tomadas de decisão e reconhecimento e comparação de padrões.

1.1.3.6 Influência da prática profissional na evolução teórica do conceito

Além da influência da Inteligência Artificial na área da classificação, a prática profissional da gestão, classificação e recuperação de informação em contexto empresarial também teve impacto na discussão teórica na área. Na conceção tradicional da Ciência da Informação (ou seja, aquela ligada às bibliotecas e arquivos), o objetivo da classificação é replicar uma realidade objetiva o mais rigorosamente possível, seguindo sempre o objetivo principal de representar o conhecimento de forma universal, exhaustiva e completa. Porém, o facto dos sistemas de classificação serem cada vez mais utilizados em contextos empresariais práticos, e em várias atividades

com finalidades distintas resultantes desses contextos específicos, fez com que o propósito principal da aplicação das classificações passasse da representação de uma realidade de forma objetiva e exaustiva, para o suporte às atividades de um determinado domínio específico e para a facilitação da comunicação entre os utilizadores, informação e sistema (Mai, Jens-Erik, 2004a). Assim, as classificações passaram de uma posição tradicional mais rigorosa e rígida na representação do conhecimento universal, para uma posição mais prática e aplicada de suporte a contextos específicos, enfraquecendo dessa maneira as fronteiras teóricas do conceito na sua conceção mais tradicional. Mai (2004a) **explica que as teorias da classificação estão hoje associadas de forma inequívoca ao aspeto da relatividade: ou seja, os sistemas de classificação aplicados às atividades profissionais não se focam na representação da realidade tal como ela existe, partindo assim do pressuposto que todas as classificações são inerentemente relativas e enviesadas – o contexto, o criador e o utilizador influenciam a classificação.** Esta relatividade da classificação dá liberdade à classificação para ser mais específica, o que acaba por ser uma mais-valia, uma vez que a torna mais útil para o contexto onde é criada e onde vai ser aplicada.

Então, **podemos dizer que a criação de uma classificação para uma organização será necessariamente enviesada pelo contexto organizacional, o que a tornará mais útil para os utilizadores – uma classificação pré-definida, mesmo refletindo a realidade do domínio do conhecimento em que uma organização se insere, não é necessariamente útil para uma organização, uma vez que a sua criação não foi influenciada pelo contexto organizacional e social específico.** Jens-Erik Mai (2010) explica que como o conhecimento e a verdade são construções sociais pensadas, produzidas e estabelecidas em interações sociais, é essencial que as classificações sejam desenvolvidas em constante colaboração com os seus utilizadores. Neste sentido, o autor afirma ainda que a credibilidade de uma classificação e a confiança dos utilizadores na mesma depende da sua capacidade de partilhar a base concetual das atividades e interações sociais do contexto onde é utilizada.

Devido aos vários domínios de aplicação das classificações, e a diferença das conceções de classificação que resultam desse facto, acaba por ser importante definir as

características do âmbito de aplicação do tipo de classificações que o método proposto neste trabalho pretende ajudar a desenvolver (e que explica a própria perspetiva em que se enquadra todo este trabalho). Mais uma vez Mai (2010) apresenta uma perspetiva lógica onde consegue enquadrar conceções distintas (conceções tradicionais e conceções mais “práticas”) dentro de uma mesma visão, apresentando três dimensões de problemas de organização de conhecimento, com âmbito e características distintas:

- ✓ **Problema grande de organização do conhecimento (Big KOP)** - exige a organização, gestão e representação de uma grande quantidade de informação para uma quantidade indeterminada de utilizadores com características, aptidões e interesses não conhecidos. Exemplos deste tipo de problema são a Internet e a maioria das bibliotecas públicas. A resolução de um problema deste tipo enquadrar-se-ia melhor numa conceção tradicional, onde o foco está na representação do conhecimento de forma universal e exaustiva;
- ✓ **Problema médio de organização do conhecimento (Medium KOP)** - requer a organização e gestão de um conjunto de informação com propósitos específicos, para ser utilizado por um conjunto de indivíduos com interesses, aptidões e conhecimentos particulares e similares, que são conhecidos por que está encarregue da resolução do problema. Este tipo enquadra-se no típico problema de organização de informação numa intranet ou portal organizacional, ou noutra tipo de contexto com características específicas. A aplicação das classificações na resolução deste tipo de problema teve o impacto na evolução teórica do conceito de “classificação” de que vimos falando nesta secção, e que suporta a perspetiva em que se enquadra este trabalho (de que falaremos mais à frente neste secção);
- ✓ **Problema pequeno de organização do conhecimento (Small KOP)** - requer a organização e gestão de um conjunto de informação de cariz individual e pessoal. Neste sentido, estes problemas são particulares de um conjunto de informação pessoal de um (ou poucos) indivíduo(s) que, como tal, é organizado pelo próprio proprietário. A resolução de um problema deste tipo não requer, obviamente, o mesmo tipo de recursos que um problema de maior dimensão, sendo normalmente resolvido de forma *ad hoc* pelo próprio indivíduo.

O método descrito e desenvolvido ao longo das páginas deste trabalho pretende resolver problemas em contextos organizacionais com características e utilizadores específicos, e, como tal, podemos concluir que este trabalho enquadra-se na resolução de problemas de dimensão média (medium KOP). Segundo Mai este tipo de problemas ainda não foram muito explorados e problematizados na literatura, possivelmente devido à especificidade de domínio, âmbito e contexto que cada problema deste tipo acarreta - o trabalho desenvolvido ao longo destas páginas tentará completar parte dessa lacuna.

Posto isto, podemos dizer que desenvolvimento de uma classificação para um contexto organizacional deve, então, ser complementado por uma análise das atividades, processos e contexto organizacional, para se perceber as interações sociais e a cultura organizacional e, assim, conseguir refletir estes aspetos na classificação. Neste sentido, a classificação “cresce” dentro da organização e é um artefacto que reflete a cultura e o contexto desse domínio – uma classificação desenvolvida exteriormente não conseguirá estabelecer esta ligação com a organização (Mai, Jens-Erik, 2004a).

É essencial que o desenvolvimento e gestão de classificações tenha em conta os requisitos e contextos organizacionais, e que este processo seja levado a cabo de forma colaborativa e transparente, de forma a refletir o verdadeiro conhecimento, natureza e cultura organizacional, a tornar claras as decisões concetuais e estruturais que estão na base da classificação, e a adquirir um carácter refletivo e interativo do contexto onde são desenvolvidas – desta maneira, a classificação será útil para a organização, credível junto dos seus utilizadores, e um instrumento que refletirá o contexto específico para onde e para o qual foi criado.

Concluindo, Jens-Erik Mai (2004a) aponta três pontos são importantes a reter no desenvolvimento e gestão de classificações:

1. O objetivo da classificação é ser uma ferramenta útil para a gestão da informação organizacional, e não representar objetivamente e rigorosamente uma realidade;
2. O processo de desenvolvimento da classificação deve basear-se também em estudos das interações sociais dos utilizadores de informação, dos seus

hábitos de trabalho e nas estruturas dos domínios que se pretendem representar;

3. Os envolvidos no processo de classificação não devem ser forçados a ser rigorosamente objetivos e neutrais – eles fazem parte do contexto e são, inerentemente, enviesados e influenciados por ele, o que faz com que a classificação também o seja. Este viés e esta influência adaptam a classificação ao contexto, o que a torna mais útil para a organização.

1.1.4 Perspetiva defendida neste trabalho e principais ideias a reter

Vistas as várias visões e conceções sobre a classificação e os aspetos relacionados com esta, é possível agora definir a perspetiva em que se enquadra o desenvolvimento deste trabalho. **Numa perspetiva da Ciência da Informação aplicada à atividade profissional, este trabalho e os seus resultados assentam na conceção da classificação como instrumento para a organização da informação num contexto específico e útil na resolução de um problema particular de organização da informação. A classificação deve assim privilegiar a sua utilidade para os utilizadores do contexto é desenvolvida e gerida, e não a representação do conhecimento de forma universal como defendem as conceções mais tradicionais.** Neste sentido, o trabalho encaixa numa visão próxima do que Mai defende (2004; 2010), ou seja, **numa perspetiva que defende a influência da prática profissional na mudança do paradigma teórico da classificação, onde esta se mostra como um instrumento de apoio às atividades e processos dentro de um contexto organizacional com interesses particulares, utilizadores específicos e características singulares (um Medium Kop, tal como Mai define), e não como uma representação do conhecimento universal.** Esta perspetiva é a que se enquadra melhor com a implementação de classificações a contextos organizacionais específicos, onde há necessidade de resolver um problema com âmbito, características e utilizadores particulares, e onde a representação universal, exaustiva e rigorosa do conhecimento não é relevante, nem eficiente, nem útil para o contexto. Contudo, os princípios e bases da conceção tradicional não deixam de ser relevantes e úteis para a construção do método - afinal de contas, na perspetiva que este trabalho defende, a classificação

continua a ser um sistema de classes que deve obedecer a uma série de regras para bem da sua coerência e funcionalidade. O aspeto essencial da visão defendida nestas páginas centra-se, sim, numa visão prática da classificação, onde o foco está na resolução de problemas de organização do conhecimento em contextos organizacionais específicos, onde existem utilizadores com interesses, aptidões e conhecimentos particulares e similares.

Com base nesta perspetiva, as ideias mais importantes a reter desta revisão da literatura científica, e que constituem os aspetos principais da base teórica em que se apoiará a construção do método, são as seguintes:

- ✓ O âmbito do contexto empresarial deve regular o processo de desenvolvimento e gestão da classificação;
- ✓ A finalidade da implementação de uma classificação num contexto empresarial é a de ser útil para os processos organizacionais desse contexto específico, e não a de representar o conhecimento de forma universal;
- ✓ Num contexto empresarial, estamos perante um problema de organização do conhecimento de dimensão média, com características específicas, e não um problema de dimensão grande onde é necessária uma representação universal e objetiva do conhecimento;
- ✓ Apesar da conceção tradicional de classificação não ser a mais indicada para contextos empresariais específicos, os seus princípios continuam a ser úteis como suporte da construção de uma estrutura coerente (devem ser interpretados como auxílio, e não como regras rigorosas);
- ✓ O processo de desenvolvimento e gestão de classificações deve ser levado a cabo de forma colaborativa e transparente, de forma a refletir as características do contexto onde foi elaborado – as classificações desenvolvidas em contextos empresariais específicos são relativas e enviesadas, e é este facto que as torna úteis para esse contexto;
- ✓ A indexação e a categorização, apesar de diferentes da classificação, são também importantes para a gestão dos conteúdos e podem-se revelar mais úteis em vários casos;

- ✓ A classificação facetada parece ser a melhor estrutura dada a sua grande flexibilidade e multidimensionalidade, no entanto cada estrutura pode-se revelar útil em determinados contextos;
- ✓ O desenvolvimento colaborativo das classificações torna a classificação credível junto dos utilizadores e motiva a sua utilização, o que beneficia a partilha de conhecimento.

1.2 Estado-de-arte tecnológico

O estado-de-arte tecnológico na área da ECM permite compreender o tipo de soluções utilizadas para a resolução de problemas relacionados com a gestão de conteúdos, o que é útil para percebermos as opções que estão disponíveis no mercado e em que medida o método a desenvolver neste trabalho poderá tirar partido das mesmas.

Dentro da indústria das tecnologias de informação, a área de ECM tem-se revelado uma das com crescimento mais acelerado nos últimos anos. O constante aumento de conteúdos criados revela-se com um dos principais impulsionadores da indústria, sempre com o objetivo de auxiliar as organizações a conseguirem controlar o caos informacional. No entanto, a inovação em ECM não se move apenas para resolver o problema do crescimento informacional – as organizações também vêm as tecnologias de ECM como soluções para potenciar a produtividade dos seus colaboradores, para promover a partilha de informação e a colaboração e, principalmente, para melhorar a eficiência e otimizar os processos organizacionais. É importante, pois, não só rever a tecnologia, mas também a componente estratégica, de forma a perceber quais opções que a indústria está a tomar na inovação da gestão de conteúdos empresariais, e ainda possíveis áreas da indústria da área da gestão da informação cuja atividade tem impacto positivo no desenvolvimento da ECM.

1.2.1 Estratégias de ECM

Apesar de ser considerada por alguns como algo inatingível, a visão de um único sistema de ECM global que consiga gerir todos os conteúdos continua a ser seguida por grande parte das organizações), sendo que muitas destas continuam, como seria de esperar, a sentir dificuldades em atingir essa meta. Na verdade, segundo um inquérito feito pela AIIM em 2011 (envolveu 650 membros da comunidade, cada um representante de uma organização de dimensão média-grande), apenas 16% das organizações haviam completado um projeto de ECM à escala global da empresa, sendo que 29% estavam ativamente a perseguir essa meta e cerca de metade tinham a

implementação desse conceito ainda em mente – estas percentagens apontam, assim, que cerca de 48% das empresas assumiam como estratégia organizacional a implementação de um sistema ECM à escala global da organização (AIIM, 2012). Apesar de este número demonstrar uma certa tendência nas estratégias de ECM, é necessário ter em conta que, numa escala universal, estas percentagens poderão tornar-se menores – basta pensar que muitas organizações não estão associadas a esta comunidade, o que dada a importância desta na área, sugere que as mesmas podem nem ter considerado um projeto de ECM desta natureza, estando mais “desprotegidas” do caos informacional.

25% das organizações respondentes ao inquérito continua a preferir estratégias que privilegiem as soluções não-globais, optando por resolver os problemas de ECM de forma departamental ou local. Esta percentagem ainda significativa pode ser explicada pelas organizações com estruturas mais descentralizadas, no entanto se considerarmos que o espaço web “elimina”, até certo ponto, as distâncias, é razoável especular que esses 25% das organizações podem sofrer, ou vir a sofrer, problemas de coesão na gestão de informação, que têm consequências negativas a nível da partilha de informação e da comunicação inter-organizacional.

Por último, é importante referir que, segundo os dados recolhidos pela AIIM, 14% das 650 organizações não possuem estratégia de ECM definida – isto é indubitavelmente negativo, uma vez que demonstra uma certa despreocupação de algumas organizações com o aumento da quantidade de conteúdos, e também um certo desconhecimento do valor da informação e do “conhecimento organizacional”. Este número poderá também estar relacionado com a falta de recursos para levar a cabo um projeto desta natureza, porém a crescente disponibilização de boas alternativas opensource em termos de tecnologias ECM poderia “contornar” essas limitações.

O estudo realizado pela AIIM indica também que o facto de mais de metade das organizações (57%) ter recorrido à personalização das aplicações departamentais, dificultou a possibilidade de “globalizar” estas mesmas no resto da organização, o que revela novamente que os requisitos específicos restringem a habilidade das organizações em aplicar ferramentas de ECM de modo eficaz e global. Este aspeto pode levar à insistência na aplicação de soluções de forma local/departamental, o que

possivelmente leva à origem de problemas na coesão entre departamentos, na partilha de informação e na adaptabilidade/mudança organizacional.

Apesar de tudo, este “sistema único”, que consiga ser a fonte de informação singular de uma organização, tem sido procurado através de diferentes estratégias – alguns projetos optam pela ligação de repositórios, outros privilegiam a integração de aplicações e ferramentas, e outros seguem ainda pela via da implementação de portais de busca. Este aspeto revela que as opções a tomar na gestão de conteúdos organizacionais dependem diretamente do contexto específico e dos requisitos de cada organização, o que prova que a tecnologia não deve ser aplicada “cegamente”.

Na próxima secção veremos algumas das tecnologias mais utilizadas para a gestão de conteúdos empresariais.

1.2.2 Tecnologias de ECM

O sistema Microsoft SharePoint é a principal referência no que diz respeito às tecnologias de ECM, sendo considerada como uma peça essencial na gestão de conteúdos por grande parte das organizações. Para corresponder aos requisitos e contextos específicos, as organizações tentam complementar o SharePoint com outras tecnologias, sejam elas repositórios de informação, aplicações interoperáveis, componentes open-source ou add-ons para a ferramenta da Microsoft. No entanto, esta última é, também, por vezes implementada não como substituto de aplicações existentes, mas sim como mais uma ferramenta acessória para funcionar “ao lado” de outras ferramentas – este aspeto demonstra a dificuldade que ainda existe em encontrar uma solução única e eficaz para o problema da gestão de conteúdos, o que leva as organizações a optarem por tentar implementar novas ferramentas para solucionar os problemas que vão surgindo, que acabam muitas vezes por não ser totalmente interoperáveis com as que já existiam (AIIM, 2012).

Apesar disso, o facto de se implementar o SharePoint como peça central do sistema ECM não se revela por si só uma solução eficaz - os resultados do inquérito da AIIM indicam que apenas cerca de 25% das implementações da ferramenta da Microsoft se aproxima de um estado otimizado de funcionamento. Isto sugere possíveis problemas na implementação da ferramenta, que poderão estar relacionados não só

com falta de recursos e esforço, mas também com dificuldades na adaptação da ferramenta aos requisitos específicos da organização, sejam elas provenientes da necessidade de desenvolvimento de classificações ou do estabelecimento de padrões para a organização dos conteúdos.

Apesar da “popularidade” da ferramenta Microsoft, existem cada vez mais soluções em termos de ferramentas e aplicações para gestão de EMC, desde alternativas opensource, até soluções desenvolvidas por grandes empresas da área da tecnologia. De seguida estão descritas algumas das soluções tecnológicas mais utilizadas (CMSCritic, 2012):

Software	Empresa	Descrição sucinta
Alfresco	Alfresco Software	• Opensource (código aberto que pode ser modificado e melhorado pelos utilizadores, o que permite que o Alfresco possa evoluir muito rapidamente) • Várias versões (versões mais avançadas requerem subscrição) (Alfresco, 2012)
Documentum	EMC Corporation	• Funções para gerir todo o tipo de conteúdos • Vários serviços como gestão de documentos, gestão de conteúdos Web, gesto de processos organizacionais, etc. • Conteúdo armazenado em vários servidores, mas disponível num ambiente unificado (CMSCritic, 2012)
FileNet	IBM	• Ferramenta da gestão de conteúdos, segurança, armazenamento, e processos organizacionais • Providencia um único repositório para conteúdo e um modelo de metainformação, assim como workflows automáticos • Integrável com o MS SharePoint e Office (IBM, 2012)
Laserfiche Avante e Laserfiche Rio	Laserfiche	• Vários recursos para digitalização, gestão de e-mail, segurança, workflow • Interoperável com software da Microsoft (Laserfiche, 2012)
Nuxeo	Nuxeo	• Plataforma opensource • Construída para ser fácil de customizar (CMSCritic, 2012)

Plone		Plone Foundation	• Plataforma opensource • Bastante utilizada na gestão de conteúdos criados no âmbito de projetos • Design e utilização simples
OpenText Suite	ECM	OpenText	• Integra várias tecnologias de gestão de informação, desde gestão documental e gestão de e-mail até gestão do ciclo de vida da informação • Possui ferramentas para captura de documentos • Integra opções colaborativas de Web 2.0, como wikis, blogs, fóruns, entre outras (OpenText, 2012)
Oracle Webcenter Content		Oracle	• Capaz de gerir vários tipos de conteúdos • Possui opções de gestão do ciclo de vida da informação • Capaz de integrar outras aplicações (CMSCritic, 2012)
SharePoint		Microsoft	• Plataforma mais utilizada do mercado • Integrável com outros tipos de sistemas de gestão organizacionais (como Enterprise Resource Planning ou Customer Relationship Management) • Utilizado normalmente como portal de intranet (CMSCritic, 2012)

Tabela 1 - Vários tipos de ferramentas para a gestão de conteúdos empresariais

Dado o crescimento da indústria, é de salientar que existem muitas mais soluções para a gestão de conteúdos empresariais além das que foram indicadas na tabela acima. Especialmente no que toca a sistemas de gestão de conteúdos Web, existe uma grande paleta de boas soluções, que permitem facilitar a gestão de um site, como por exemplo Concrete5, CMS Made Simple, Liferay, ProcessWire, Texpattern entre outras soluções que receberam críticas positivas nos últimos tempos (Johnston, 2012). No entanto, estes sistemas, pelo facto de serem focados na gestão de conteúdos num ambiente de website, carecem de várias opções que os impossibilita de serem implementados em organizações de média-grande dimensão, tais como gestão de documentos, captura de informação, gestão de e-mail, grandes repositórios de conteúdos, entre outras.

No suporte à classificação dos conteúdos empresariais neste tipo de sistemas, são geralmente utilizadas taxonomias corporativas, classificações baseadas em mapas de

conceitos, ou quando é necessário uma estrutura mais avançada e “inteligente”, opta-se mesmo pelo desenvolvimento de ontologias.

As **taxonomias corporativas** utilizadas pelas organizações são, basicamente, a aplicação de uma estrutura em hierarquia a um contexto corporativo, na tentativa de construir uma arquitetura de informação que conecte diferentes repositórios de forma à recuperação poder ser feita através de um único ponto central. Uma das funções principais é também providenciar um sistema de navegação estruturado para coleções de conteúdos. Estas taxonomias podem incluir conteúdos retirados de tesauros e ontologias (Delphi Group, 2004).

Em termos de **ontologias**, já caracterizadas e discutidas anteriormente neste trabalho, existem vários softwares para a sua construção (Protégé, um dos principais), no entanto envolvem a utilização de linguagens de programação e de lógica formal e estão mais orientados para uma utilização na inteligência artificial e engenharia de sistemas, o que não se mostra o mais indicado para a construção do método acessível e ágil que se pretende neste trabalho.

Os **mapas de conceitos** são estruturas gráficas que permitem organizar e representar conhecimento (Novak e Cañas, 2006). São constituídas por conceitos, representados através de caixas ou círculos, e por relações entre estes indicadas através de uma seta e de uma descrição textual breve e básica especificando o teor da relação. Esta estrutura permite que sejam construídas proposições incluindo dois ou mais conceitos e utilizando as relações entre estes. São organizados num tipo de hierarquia posicional, com o conceito mais geral no topo e os mais específicos no fundo. Além disso, podem incluir links para outros mapas de conceitos de domínios diferentes, mostrando as relações existentes entre determinados conceitos de contextos distintos. Têm, assim, uma utilização bastante simples e amigável para o utilizador, sendo por isso muito utilizadas colaborativamente como base para futuras classificações. Alguns dos softwares mais interessantes para o desenvolvimento deste tipo de estruturas são:

- IHMC CmapTools – é uma ferramenta em vários contextos (escolas, instituições governamentais, organizações, projeto, etc.) que permite que os utilizadores construam os mapas de conceitos no seu computador pessoal, partilhem a sua criação com outros utilizadores através da

internet, liguem os seus mapas com outros mapas, editem mapas sincronizadamente com outros utilizadores, entre outras possibilidades (IHMC CmapTools, 2013);

- ConceptME – é uma ferramenta que permite que os utilizadores possam recorrer a várias ajudas para a criação do seu mapa, tais como extração terminológica, templates e contextos. Outra das principais inovações desta ferramenta é a funcionalidade de Merge, que permite a deteção de conceitos equivalentes em modelos diferentes através da comparação do nome, o que melhora significativamente a utilização deste tipo de tecnologia em ambiente colaborativo (Costa et al., 2012)

O apanhado das tecnologias feito nesta secção permitiu ter uma visão geral do tipo de opções em termos de software para a gestão de conteúdos. De seguida, serão revistas algumas das áreas diretamente relacionadas com a ECM, e que apresentam desenvolvimentos importantes para a gestão de conteúdos empresariais.

1.2.3 Business Intelligence e Enterprise Information Management

A área de Business Intelligence (BI) também tem merecido atenções por parte das organizações, particularmente desde que estas perceberam o valor e importância que a informação poderá ter no seu futuro. Existem várias definições e abordagens para o conceito de BI, porém parece existir um consenso de que este designa o processo de utilizar e gerir a informação de forma a suportar as tomadas de decisão nas organizações (Rouhani, Asgari e Murhosseini, 2012). Neste sentido, as tecnologias de BI - benchmarking, data mining, análise estatística, entre muitas outras - fornecem funcionalidades de recolha, análise e acesso à informação dos processos organizacionais, de forma a ser possível detetar tendências e prever cenários futuros tendo em conta o histórico e estado atual do contexto empresarial – o objetivo principal da BI é, assim, ajudar a organização a tomar as melhores decisões possíveis (Srivastava, 2012). De certa forma, a BI permite que as organizações aprendam com a sua informação e desenvolvam o seu próprio conhecimento, no sentido de tomarem decisões ponderadas e informadas.

As principais empresas envolvidas na área da ECM também disponibilizam, naturalmente, soluções de BI, o que faz com que este mercado também seja bastante competitivo: a Oracle disponibiliza várias aplicações para análise de dados, a IBM tem a sua ferramenta Cognos com várias opções dinâmicas e intuitivas de análise e entrega de outputs, a Microsoft fornece as suas soluções focadas na monitorização e planeamento e compatíveis com as restantes aplicações da empresa, e vários outros competidores, como a QlickTech, a JasperSoft e Pentaho, apresentam no mercado as suas próprias tecnologias, o que permite que os vários tipos de clientes possam escolher a tecnologia mais alinhada com as suas necessidades e características (Rudy, 2012). As previsões indicam que o mercado de BI vai continuar a crescer e a modificar-se com novas tendências tecnológicas. A Forrester (Evelson, 2011) prevê que nos anos vindouros a indústria de BI integre cada vez mais as inovações da web 2.0., fornecendo mais controlo ao utilizador, aproveitando as funcionalidades da computação em nuvem ou “cloud” (disponibilizar os recursos tecnológicos como um serviço através da web), e apostando cada vez mais na mobilidade que as novas tecnologias tornam possível.

Nos últimos anos, o conceito *Enterprise Information Management* tem surgido como o próximo passo no futuro da gestão da informação organizacional. Apesar de não existir uma definição consensual do conceito, podemos dizer que, de forma simples, EIM conjuga as abordagens à gestão da informação da ECM e da BI, e adota uma visão única para toda a informação, sempre numa perspetiva empresarial (T-systems, 2012). Ou seja, enquanto a BI e a ECM se focam de certa maneira numa “parte” da informação organizacional – informação estruturada e não estruturada, respetivamente -, a EIM não faz qualquer distinção entre tipos de informação, “considerando” assim a gestão de toda a informação organizacional. Desta forma, a EIM pretende tornar a informação não-estruturada em informação estruturada e a informação estruturada em conhecimento, com a ajuda de várias soluções tecnológicas, e assim permitir que as empresas possam otimizar a sua gestão da informação de forma a melhorar a tomada de decisão e suportar os processos organizacionais. Esta visão integrada da informação organizacional, e do sistema de informação, por consequência, aproxima-se da visão da Ciência da Informação sobre o sistema de informação integrado e, também, combina com o relativamente novo conceito de Enterprise Search, que pretende tornar pesquisável e recuperável toda a informação organizacional a partir de um único sistema (de que falaremos na secção seguinte).

1.2.4 Enterprise search

O crescimento da quantidade de conteúdos empresariais tem “forçado” as organizações a implementar vários repositórios de informação em paralelo - intranet, caixa de e-mail, bases de dados -, de forma a ter espaço para armazenar toda essa massa informacional. Isto origina, como seria de esperar, problemas na recuperação de informação, uma vez que as informações ficam armazenadas em vários tipos de repositórios e podem ter sido indexados de formas diferentes, o que torna difícil a missão dos motores de busca. Em resposta a esta situação, a indústria tem apostado, nos últimos anos, no desenvolvimento de tecnologias que tornem pesquisável e recuperável a informação armazenada em múltiplos e diferentes repositórios – este tipo de software é agora conhecido como *enterprise search*. Esta tecnologia permite, então, pesquisar toda a informação organizacional a partir de um mesmo motor de busca, semelhante ao que acontece com os típicos motores de busca da web. Para atingir este objetivo, este tipo de software possui capacidades que lhe permitem indexar os conteúdos e comparar o resultado deste processo com as expressões de pesquisa introduzidas pelos utilizadores, no sentido de recuperar, teoricamente, os resultados mais relevantes para a “pergunta” introduzida – para indexar os conteúdos, o software necessita de coletar os vários tipos de conteúdos dos vários repositórios e, de seguida, processar e analisar o seu conteúdo, utilizando técnicas habituais neste tipo de processo, como por exemplo a frequência do termo. Dada a variedade de conteúdos empresariais, é necessário que estes sistemas de pesquisa organizacionais possuam várias funcionalidades que permitam aos utilizadores limitar a expressão de pesquisa, de forma a possibilitar uma maior precisão nos conteúdos recuperados – estas funcionalidades passam por pesquisas facetadas, sistemas de tagging na captura de conteúdos, *clustering* (agrupar conteúdos semelhantes), entre outros (Regli e Bloem, 2010). Além disso, dada a necessidade de segurança na informação organizacional, este tipo de sistemas possui funcionalidades de controlo de acesso aos conteúdos, o que permite restringir o acesso conforme o tipo de utilizador.

As tecnologias de *enterprise search* são, assim, complementadas com ferramentas direcionadas para a indexação automática de conteúdo no momento da sua captura – no entanto, dada a subjetividade inerente à análise do conteúdo a indexar, e a falta de capacidade interpretativa das tecnologias, é razoável afirmar que

técnicas automáticas de indexação não serão exatamente precisas em todo o conteúdo. Apesar dessa desvantagem, o facto das mesmas serem muito mais rápidas que as técnicas humanas de indexação faz com que sejam imprescindíveis para as empresas, na tentativa de “organizar” o crescente caos informacional.

Em termos de soluções tecnológicas propriamente ditas, são várias as empresas que vêm desenvolvendo software para *enterprise search*, desde a IBM, e a Oracle, até à Autonomy, a Endeca e a Coveo. Na tabela seguinte estão descritas algumas das principais soluções (de acordo com os estudos da Forrester Research):

Software	Descrição sucinta
Autonomy	• Boa arquitetura tecnológica • Boas funcionalidades de segurança • Ferramenta mais complexa e dispendiosa
Coveo	• Possui uma interface baseada na web, com suporte a várias linguagens • Software relativamente mais barato
Endeca	• Foco na pesquisa de bases de dados, XML, RSS •
Fast (da Microsoft)	• Componente nuclear do SharePoint, mas pode ser adquirido como um produto único • Indicada para empresas que apostam nas plataformas da Microsoft
Google Enterprise Search	• Fácil e rápido de implementar • Software orientado para ser conveniente para o cliente • Funcionalidades de segurança limitadas
InQuira	• Funcionalidades avançadas para suporte da linguagem natural, utilizando vários dicionários personalizáveis
OmniFind (da IBM)	• Limitado a nível de funcionalidades de gestão de taxonomias
Recommind	• Orientado para empresas da área legal
Secure Enterprise Search (da Oracle)	• Limitado a nível de funcionalidades de gestão de taxonomias
Vivisimo	• Interface e ferramentas sociais muito flexíveis

Tabela 2 - Alguns dos softwares mais populares da área de Enterprise Search

1.2.5 Enterprise 2.0

Um dos aspetos a realçar do estudo da indústria feito pela AIIM é o facto das atividades e processos colaborativos estarem a ser cada vez mais vistos como um aspeto essencial para a ECM – mais de metade das organizações respondentes ao inquérito afirmaram utilizar ferramentas colaborativas dentro da empresa, suportadas não só pelo SharePoint mas também por outras plataformas (AIIM, 2012). Esta tendência mostra que as inovações da web 2.0 estão a ser utilizadas pela indústria das tecnologias da informação e comunicação, permitindo assim que os colaboradores das organizações, e as próprias organizações, possam interagir e partilhar informação através da web, e participar e contribuir em atividades colaborativas.

A esta nova tendência da indústria foi atribuído o termo “enterprise 2.0”, por Andrew McAfee. Enterprise 2.0. designa, então, a utilização das novas e emergentes plataformas e tecnologias sociais da web 2.0 nas empresas, permitindo que as pessoas possam interagir e colaborar através da web, e formar comunidades virtuais (McAfee, 2006). Esta tendência foca-se, acima de tudo, em dar mais poder aos colaboradores, promovendo a interação e colaboração, e possibilitando que os utilizadores possam assumir um papel mais ativo e influenciador na organização. Assim, os colaboradores possuem mais liberdade para criar e indexar os seus próprios conteúdos, promovendo uma cultura de colaboração que impacta positivamente a gestão da informação e o envolvimento dos colaboradores numa organização.

Algumas das tecnologias da web 2.0. que têm sido adaptadas ao ambiente empresarial são:

- ✓ **Blogging**
- ✓ **Wikis** – permitem a partilha de informação num ambiente colaborativo e amigável;
- ✓ **Plataformas para comunidades empresariais** - plataformas que permitem a criação de um ambiente virtual de partilha de opiniões ou informação, como por exemplo um fórum.
- ✓ **Computação em nuvem ou “cloud computing”** – disponibilização de recursos computacionais em forma de serviço através da rede Web. Os

recursos podem ser software, plataformas, infraestruturas, redes, bases de dados, repositórios, entre muitos outros.

- ✓ **Tagging** – categorização o de conteúdos feita pelos próprios utilizadores, criando folksonomias.

1.2.6 Ideias a reter quanto ao estado-de-arte da gestão de conteúdos

A revisão do estado-de-arte tecnológico permitiu alcançar algumas conclusões importantes:

- ✓ As organizações perseguem a estratégia de “sistema único” e global para a gestão dos conteúdos, no entanto continuam a sentir dificuldades na integração das plataformas. A coesão e normalização na gestão da informação são vistas como muito importantes pela maioria das organizações;
- ✓ As opções a tomar na gestão de conteúdos organizacionais dependem diretamente do contexto específico e dos requisitos de cada organização, o que prova que a tecnologia não deve ser aplicada “cegamente”;
- ✓ A variedade de soluções em termos de sistemas de gestão de conteúdos é elevada, sendo que as organizações suportam esse tipo de tecnologias através da construção de taxonomias corporativas, de ontologias e, também, de mapas de conceitos;
- ✓ As áreas de Business Intelligence e de Enterprise Information Management fornecem soluções que pretendem valorizar o peso da informação no momento da tomada de decisão, e também a integração de toda a informação num só conjunto, duas perspetivas que se aproximam da defendida pela área da Ciência da Informação. Como é evidente, isto mostra que a indústria está atenta à crescente importância da informação no contexto organizacional, e também ao conhecimento proveniente da área de CI;
- ✓ A Enterprise Search é mais uma área que privilegia a integração dos sistemas e repositórios de informação, de forma a facilitar o trabalho do utilizador no momento da pesquisa. Assim, as soluções tecnológicas no

âmbito desta área demonstram uma valorização do papel do utilizador, ao procurar formas de melhorar a interação tecnologia-utilizador, e também uma preocupação em organizar o “caos” informacional;

- ✓ As soluções da área da Enterprise 2.0 demonstram a importância das atividades e processos colaborativos na gestão da informação, focando a utilização das novas e emergentes plataformas e tecnologias sociais da web 2.0 nas empresas.

2. Metodologia de investigação

Tendo em conta que o principal objetivo deste trabalho é a criação de um artefacto – um método para o desenvolvimento e gestão de classificações à medida –, é necessário que a metodologia utilizada se adeque a esta meta. Como tal, a seleção da metodologia teve em conta alguns critérios que decorrem diretamente da natureza do tema e dos resultados esperados, nomeadamente:

- ✓ Existência um problema prático, “real” e com características e contexto complexos – como desenvolver e gerir eficazmente uma classificação à medida de determinado contexto específico?
- ✓ Problema pouco estudado na literatura e falta de soluções práticas concretas motivam necessidade de solução inovadora e útil – investigação exploratória, sem modelos de partida, com o intuito de chegar a uma conclusão satisfatória
- ✓ Criação de um artefacto prático para resolver o problema – método eficaz para o desenvolvimento e gestão de classificação à medida

Tendo em conta estes aspetos, selecionou-se **a abordagem metodológica de *design science* aplicada à investigação em sistemas de informação**, uma vez que parece ser o método que mais garantias fornece para atingir os objetivos e para estudar o problema, tal como é possível ver de seguida na descrição da metodologia.

2.1 Design science

A design science baseia-se nos métodos e metodologias de investigação das ciências aplicadas e da engenharia, tendo, desta forma, um cariz muito prático, onde o objetivo reside, sobretudo, na resolução de problemas concretos e reais. Na área da ciência da informação, esta abordagem metodológica envolve o desenvolvimento de artefactos inovadores e a análise e avaliação do seu uso, com a finalidade de se obterem soluções eficazes para problemas concretos encontrados. Os artefactos criados são diversos, podendo ser modelos, métodos, práticas, linguagens, aplicações, entre muitos outros. O carácter inovador destes torna-os preponderantes na resolução de problemas encontrados em diversos contextos, especialmente em organizações – é, devido a este

motivo, uma ciência que têm intervenção direta (e, se bem sucedida, benéfica) no mundo, através da inovação.

Os problemas que estão tipicamente na origem de uma investigação modelada pela design science têm características próprias e complexas: os requisitos são instáveis e os contextos pouco definidos, as interações e relações entre os vários elementos do problema e a possível solução são complexas e, também, a possível chegada a uma solução para o problema está dependente da criatividade e de outras capacidades cognitivas do investigador. Tendo em conta a falta de literatura sobre métodos de desenvolvimento de classificações e a divergência de opiniões e perspetivas nesta área, podemos concluir que o problema em estudo neste trabalho apresenta características que o tornam um bom “candidato” a ser investigado segundo uma abordagem metodológica baseada na design science – e por aqui passa também uma das razões da escolha desta metodologia.

Numa abordagem metodológica de design science típica, pretende-se que o investigador, durante o processo de investigação, compreenda melhor o problema que está na origem da investigação, o que lhe permitirá obter soluções mais eficazes e úteis. Este artefacto permitirá resolver problemas que até aí não tinham solução, ou então, trará formas mais eficientes e eficazes de resolver problemas – de uma forma ou de outra, o resultado de uma investigação baseada nesta abordagem é único e inovador, provocando impacto no mundo. A abordagem metodológica da *design science* adota, assim, uma posição ativa e interventiva no contexto que investiga, na tentativa de compreender e interpretar melhor o problema, que permite chegar a uma solução eficaz (Hevner et al., 2004).

A ilustração 5 ilustra um dos vários frameworks existentes para a abordagem metodológica de design science (Vaishnavj e Kuechle, 2004). Apesar da literatura ser bastante rica no que diz respeito a frameworks dos processos de design, selecionou-se este porque é bastante simples e compacto, e dá o ênfase necessário às fases do processo, à criação de conhecimento novo durante este, e o output típico de cada fase.

- ✓ A primeira fase (**Awareness of problem** – tomar consciência do problema) consiste na percepção da existência de um problema real. O output desta fase é uma proposta de investigação;

- ✓ A segunda fase (**Suggestion** – sugestão) tem como grande meta a

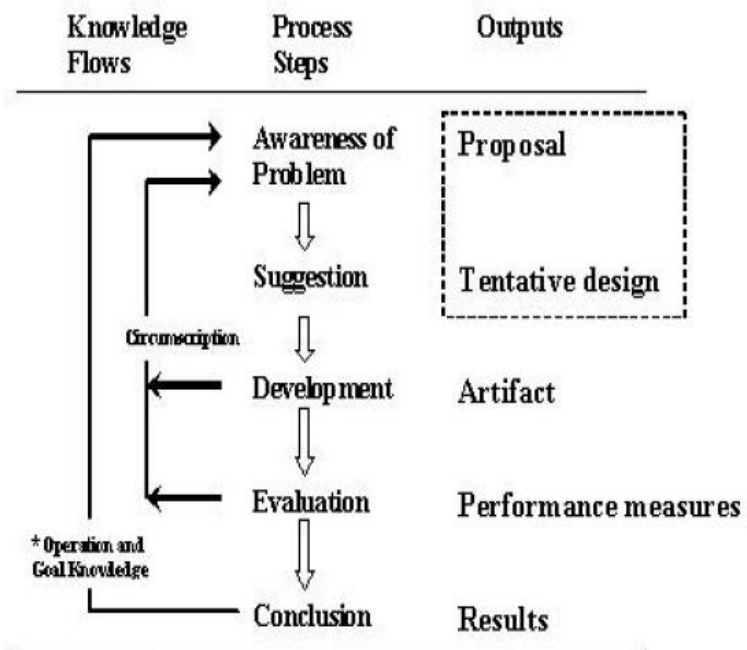


Ilustração 5 - Fases do processo de design science

- sugestão de uma possível solução para o problema. Como tal, esta fase será a que requererá mais esforço criativo por parte do investigador, na tentativa de criar um artefacto inovador a partir de informações, artefactos e elementos já existentes. O output desta fase é a própria sugestão de artefacto, e está interligado com a proposta de investigação da primeira fase, uma vez que, normalmente, esta proposta já conta com algum tipo de sugestão de resolução do problema;
- ✓ A terceira fase (**Development** – desenvolvimento) é a fase de desenvolvimento (e implementação) da solução sugerida na fase anterior, ou seja, a construção do artefacto. A duração e as tarefas desta fase variam muito consoante o tipo de artefacto a desenvolver (um modelo ou método não acarreta o trabalho tecnológico requerido na implementação de um algoritmo ou sistema informático). O output desta fase é, como óbvio, o artefacto.
 - ✓ A quarta fase (**Evaluation** – avaliação) consiste na avaliação do artefacto construído na fase anterior, de acordo com uma série de critérios implícitos ou explícitos. Os desvios em relação ao esperado e os problemas verificados devem ser registados e analisados, e devem ser sugeridas hipóteses para os mesmos. O conhecimento tirado nesta fase será essencial caso seja necessário reformular e sugerir uma nova solução – como tal, este processo não é linear, mas sim um

processo cíclico de tentativa-erro e aprendizagem até alcançarmos uma solução eficaz. O output desta fase é um relatório de avaliação do artefacto, contendo a informação necessária para se perceber se o artefacto é uma solução suficiente para o problema.

- ✓ A quinta fase (**Conclusion** – conclusão) é, naturalmente, a última etapa do processo. Esta fase acontece quando o artefacto que se avaliou é suficientemente eficaz para resolver o problema percebido na primeira fase. Ou seja, mesmo que o artefacto possua desvios em relação às expectativas, os resultados obtidos podem ser suficientemente bons para resolverem o problema – ou seja, o artefacto é uma solução suficiente para o problema. Nesta fase, é importante consolidar todos os resultados, mas também reconhecer e registar todo o conhecimento “absorvido” durante a investigação quanto possível, bem como registar todos os problemas e anomalias que não conseguiram ser explicadas (funcionarão como um possível objeto de investigações futuras). O output desta fase serão os resultados globais da investigação.

Durante todo este processo, o investigador deverá ter uma postura de aprendizagem, uma vez que a sugestão de solução e desenvolvimento e avaliação do artefacto permitem que ele possa adquirir conhecimento e compreender melhor o problema que pretende resolver. Neste sentido, os fluxos de conhecimento são inerentes a este processo cíclico de tentativa-erro, e permitem que o investigador compreenda cada vez melhor o contexto e a natureza do problema, o que o ajudará a sugerir progressivamente melhores soluções.

2.2 Adaptação da metodologia aos objetivos da dissertação

Vistos os pontos gerais da abordagem metodológica da design science, é necessário agora adaptar a mesma às características e objetivos desta dissertação:

- ✓ **Tomar consciência do problema** – esta fase envolveu à elaboração proposta e ao projeto de investigação, onde foi possível definir, delimitar e compreender de melhor forma o problema a resolver. Nesta fase foram levadas a cabo as seguintes atividades:
 - *Pesquisa e análise de literatura científica sobre o tema;*
 - *Pesquisa e análise de informação sobre o estado-de-arte;*

- *Elaboração da proposta e do projeto de dissertação;*
- ✓ **Sugestão** – nesta dissertação, as fronteiras entre o término da primeira fase e o início desta são de difícil delimitação, uma vez que o problema a resolver envolve várias perspetivas que precisam de exploração. Neste sentido, durante a fase criativa de sugestão da solução, continuar-se-á a compreender melhor o problema a resolver e a explorar as várias perspetivas em confronto no âmbito da classificação e da gestão de conteúdos. Nesta fase, é importante referir que se seguiu uma abordagem exploratória mais explicativa, uma vez que a natureza do trabalho obrigava a um maior foco nas questões teóricas e concetuais, de forma a se conseguir uma maior adequação entre a prática profissional e o cariz científico da Ciência da Informação e da Representação do Conhecimento. Tendo em conta que a *design science* nasce da identificação de um problema real, e que a ciência explicativa se incide na explicação de fenómenos reais, decidiu-se que uma abordagem mais “explicativa” nesta fase seria interessante, de forma a explorar mais aprofundada e exaustivamente o problema e as questões concetuais e teóricas envolvidas no contexto. Além disso, o facto da perspetiva em que se insere este trabalho privilegiar a aplicação prática de classificações a contextos específicos, dá ao trabalho uma vertente bastante prática que exige a recolha de informação qualitativa em contextos profissionais reais. Neste sentido, nesta fase essa recolha de informação será fundamental para perceber os contextos profissionais, de forma a perceber de que forma o modelo deverá ser estruturado - esta fase privilegia assim uma complementaridade entre a revisão teórica e a informação recolhida nos contextos práticos e profissionais. Posto isto, esta fase abarcou as seguintes atividades:
 - *Análise de literatura científica;*
 - *Realização de entrevistas exploratórias a vários profissionais com ligações à classificação e gestão de conteúdos em projetos (colaboradores da instituição INESC Porto), para perceber os vários contextos, processos, problemas, recomendações e outros aspetos inerentes à classificação e gestão de conteúdos, de forma a ser possível encontrar e deduzir padrões na gestão de conteúdos num contexto profissional complexo;*

- *Análise de portais para gestão de conteúdos no âmbito de diversos projetos;*
- *Definição dos aspetos e características do artefacto a desenvolver (método para o desenvolvimento e gestão de classificações à medida);*
- ✓ **Desenvolvimento** – nesta fase, foi desenvolvido o artefacto cujas características e aspetos gerais haviam sido delimitados no final da fase anterior. Nesta dissertação, o artefacto a ser criado enquadra-se no conceito geral de “método”. Um método, segundo March e Smith (1995) é uma forma de levar a cabo atividades orientadas por objetivos – um conjunto de passos, como uma *guideline*, utilizado na realização de uma tarefa ou atividade. Neste sentido, nesta fase realizou-se a seguinte atividade:
 - *Desenvolvimento do método;*
- ✓ **Avaliação** – nesta fase, foi avaliado o artefacto desenvolvido na fase anterior, através da aplicação do mesmo num contexto profissional (serviço de consultoria do INESC Porto) e do registo dos aspetos e problemas experienciados nessa avaliação. A avaliação do método tem como foco os seguintes critérios:
 - Funcionalidade (capacidade do método para alcançar resultados)
 - Eficiência (relação tempo de aplicação vs complexidade do resultado)
 - Usabilidade (facilidade de aplicação)
 - Utilidade (eficácia do resultado do método para satisfazer necessidades)

Esta fase englobou as seguintes atividades:

- *Aplicação do método num contexto profissional complexo e avaliação da performance do mesmo;*
- *Registo dos acontecimentos e problemas durante a avaliação;*
- *Possíveis reformulações ou ajustes do método, e nova avaliação (que requerem passagem pela segunda e terceira fase novamente);*
- ✓ **Conclusão** – esta fase traduz-se na redação final da dissertação, com a apresentação e explicação de todo o trabalho realizado, de todo o conhecimento científico criado, e com as conclusões finais sobre todo o processo de investigação focando, naturalmente, o método criado. Assim, podemos definir a seguinte atividade para esta fase:

- *Redação do relatório final de dissertação.*

O processo de investigação que suportou esta investigação teve em conta uma série de princípios (Hevner et al. 2004), que permitiram que o trabalho realizado seguisse dentro dos padrões da abordagem metodológica da design science, respeitasse as suas boas práticas, e que a escolha da abordagem fosse adequada à natureza do problema em estudo nesta dissertação:

1. **Criação de um artefacto** – o resultado da investigação é um artefacto inovador e viável, criado diretamente para solucionar um problema (ou para apresentar uma melhor solução que as existentes) e para intervir no mundo;
2. **Problema relevante** – o objetivo da investigação é adquirir conhecimento sobre e compreender um problema importante e ainda não resolvido, de forma a ser possível chegar a uma solução adequada;
3. **Avaliação da solução** – o artefacto desenvolvido deve ser avaliado de forma a ser demonstrada e justificada a sua utilidade, qualidade e eficácia na resolução do problema. A solução será eficaz e completa quando for suficiente para satisfazer os requisitos do problema que pretende resolver;
4. **Contribuições da investigação** – a investigação deve providenciar contribuições teóricas, práticas e metodológicas;
5. **Rigor na investigação** – a investigação requer a aplicação de métodos rigorosos na construção e na avaliação do artefacto, especialmente no que diz respeito à aplicabilidade e generalização do artefacto;
6. **Construção do artefacto como um processo de pesquisa** – a construção do artefacto é um processo de pesquisa até à descoberta de uma solução suficientemente eficaz para o problema. A importância está em justificar que a solução encontrada resolve o problema;
7. **Comunicação da investigação** – a investigação deve ser comunicada de forma a ser compreendida por vários tipos de audiências, desde profissionais a gestores, passando por audiências mais orientadas para a tecnologia. Desta forma, as mais-valias da investigação terão maior impacto.

3. Padrões de organização e classificação de conteúdos

A perspetiva que este trabalho defende está, por natureza, ligada à prática profissional e à aplicação das classificações em contextos empresariais específicos, assim como o método que se pretende desenvolver, onde o foco está na aplicação prática. Este aspeto faz com que seja essencial perceber os processos e tarefas ligadas à classificação de conteúdos que são levadas a cabo em contextos empresariais reais, de forma a encontrar a melhor relação e complementaridade entre o conhecimento científico-teórico e a aplicação prática empresarial. Tendo em conta esta necessidade, a metodologia de investigação definida para este trabalho (que foi explicada no último capítulo) incluiu a recolha de informação, através de entrevistas de cariz exploratório, junto de profissionais envolvidos em vários projetos, e com conhecimento sobre os aspetos ligados à gestão e classificação de conteúdos realizada no âmbito desses mesmos projetos. Assim, neste capítulo essa informação será analisada e problematizada, no sentido de se estabelecerem padrões de organização e classificação de conteúdos que serão vitais para a construção do método – as características de cada contexto ajudarão a perceber quais os melhores modelos estruturais de classificação para cada tipo de contexto, definindo-te assim uma série de padrões de organização e classificação de conteúdos adequados a situações empresariais específicas, sendo possível chegar a um método com base no sistema concetual teórico, mas com o foco na aplicação prática.

As entrevistas foram realizadas com três profissionais colaboradores do INESC Porto, um instituto cuja atividade foca fundamentalmente a investigação científica e o desenvolvimento tecnológico, funcionando com um elo de ligação entre o mundo académico e o mundo empresarial da indústria e dos serviços, especialmente na área da tecnologia. Os profissionais do INESC Porto beneficiam de uma experiência de colaboração em vários projetos internacionais, tendo por isso um conhecimento prático da gestão de conteúdos em diversos contextos, e consequentemente sendo capazes de transmitir informação extremamente valiosa para o desenvolvimento do método. O cariz das entrevistas foi totalmente exploratório para que os entrevistados tivessem

mais liberdade para expressarem e transmitirem as suas ideias, experiências e preocupações. A informação recolhida das entrevistas foi analisada, e as ideias mais importantes para os objetivos deste trabalho serão interpretadas, de forma a se poder chegar a algumas conclusões quanto à aplicação prática das bases teóricas da classificação e dos padrões de organização dos conteúdos a cada tipo de contexto. Além disso, também foram analisados dois sistemas de gestão de conteúdos utilizados em dois projetos em que o INESC Porto participou, nomeadamente o portal NetChallenge e o HKnow. As informações recolhidas dos portais permitiu complementar aquela recolhida nas entrevistas, sendo uma segunda perspetiva, mais direta, onde foi possível ver na prática algumas das opções dos profissionais para a gestão dos conteúdos no âmbito de projetos. Para colmatar toda esta informação recolhida no contexto profissional prático será utilizado também o conhecimento obtido através da revisão de literatura, já problematizada neste trabalho, e também alguma da informação obtida junto do serviço de consultoria do INESC Porto, que se disponibilizou para aplicar e avaliar o método desenvolvido neste trabalho.

Antes disso, é importante perceber como é feita a gestão de conteúdos no INESC Porto, de forma geral. Neste contexto, são utilizadas, principalmente, duas plataformas para a gestão de conteúdos, nomeadamente o *Plone* e a *Dropbox*. O primeiro é um sistema de gestão de conteúdos bastante básico, com funcionalidade de criação e organização de pastas, criação e submissão de conteúdos e atribuição de categorias a estes; o segundo, apesar de não ser um sistema de gestão de conteúdos na sua essência, é utilizado para tal pelos colaboradores do INESC Porto, além da sua habitual função de sincronização de conteúdos através da utilização da *cloud*. A utilização de duas plataformas distintas significa que cada uma possui características e funções próprias dentro do contexto, nomeadamente:

- ✓ ***Plone*** – repositório das versões finais dos documentos importantes para os objetivos do projeto (os “*deliverables*”), e portal web para acesso a conteúdo sobre o projeto. A dificuldade na gestão de versões de documentos e a pouca usabilidade na submissão de documentos dificultam o uso do Plone como sistema de gestão de conteúdos durante o projeto;

- ✓ **Dropbox** – apesar de esta ferramenta não ser por definição um sistema de gestão de conteúdos, ele é efetivamente utilizado no INESC Porto como um sistema para gerir os conteúdos criados e partilhados durante o decorrer de um projeto, uma vez que permite de forma muito intuitiva através de um ambiente familiar (Windows Explorer), a partilha de versões de documentos criados através das mais-valias do *cloud computing*. A Dropbox gere, então, a informação de todos os projetos no mesmo “espaço”, permitindo também assim a gestão de conhecimento.

Veremos de seguida a informação recolhida acerca dos padrões de organização e classificação de conteúdos em cada um dos sistemas (para efeitos de trabalho, o termo “sistema de gestão de conteúdos” também será utilizado para se referir à Dropbox, uma vez que efetivamente tem essas funções no INESC Porto). No que diz respeito aos padrões de organização e classificação de conteúdos, as questões mais importantes neste âmbito estarão relacionadas com a estruturação dos sistemas de classes, com a definição dos conceitos e termos a incluir nestes sistemas, e, não menos importante, com as formas possíveis de adaptação das classificações aos sistemas de gestão de conteúdos. Será analisada, então, a informação recolhida relevante para cada um destes aspetos, de forma a ser possível definir as características essenciais para a construção do método.

3.1 Estruturação da classificação

O primeiro aspeto a retirar da informação obtida nas entrevistas é o facto de o processo de classificação não ser normalmente realizado tal como é definido “classicamente” – nos vários projetos onde os entrevistados estiveram envolvidos, o processo de classificação foi feito indiretamente, ou seja, não passou pela atribuição direta de uma “classe” a cada conteúdo. Em vez disso, o processo de classificação envolvia tipicamente o armazenamento de cada conteúdo numa “pasta”, integrada numa estrutura hierárquica de pastas.

“Nós fazemos isso através de uma plataforma, que por vezes é um repositório de ficheiros, e que está organizado por pastas correspondentes às tarefas ou aos resultados de projetos. Os documentos vão sendo inseridos nessas pastas.”

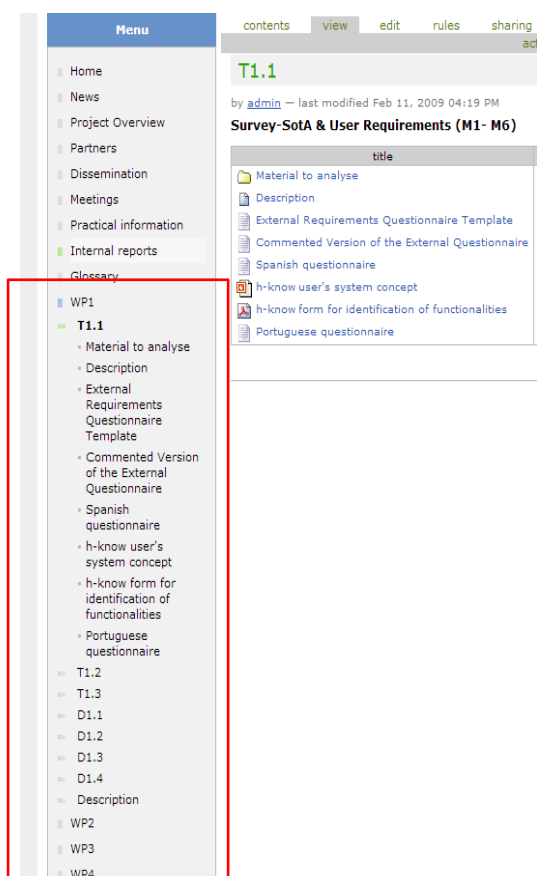
Anónimo 1

Como é possível perceber pela afirmação do entrevistado, as funcionalidades de atribuição de categorias existente nos sistemas de gestão de conteúdos não foram utilizadas para classificar os conteúdos, assim como não foi definida qualquer estrutura hierárquica de classes para conduzir a atribuição de possíveis categorias aos conteúdos. Podemos afirmar então que a estrutura hierárquica de pastas - em que cada pasta pode ser considerada como uma classe - é a estrutura de classificação normalmente utilizada no tipo de projetos que o INESC Porto participa.

Os critérios que estão na base da criação e gestão da estrutura de pastas estão normalmente relacionados com a própria organização das atividades do projeto:

“[...] De certa forma, a maioria dos projetos acaba por ter uma estrutura mais ou menos normalizada, que normalmente está organizada de acordo com a proposta de projeto que nos chega.”

Anónimo 3

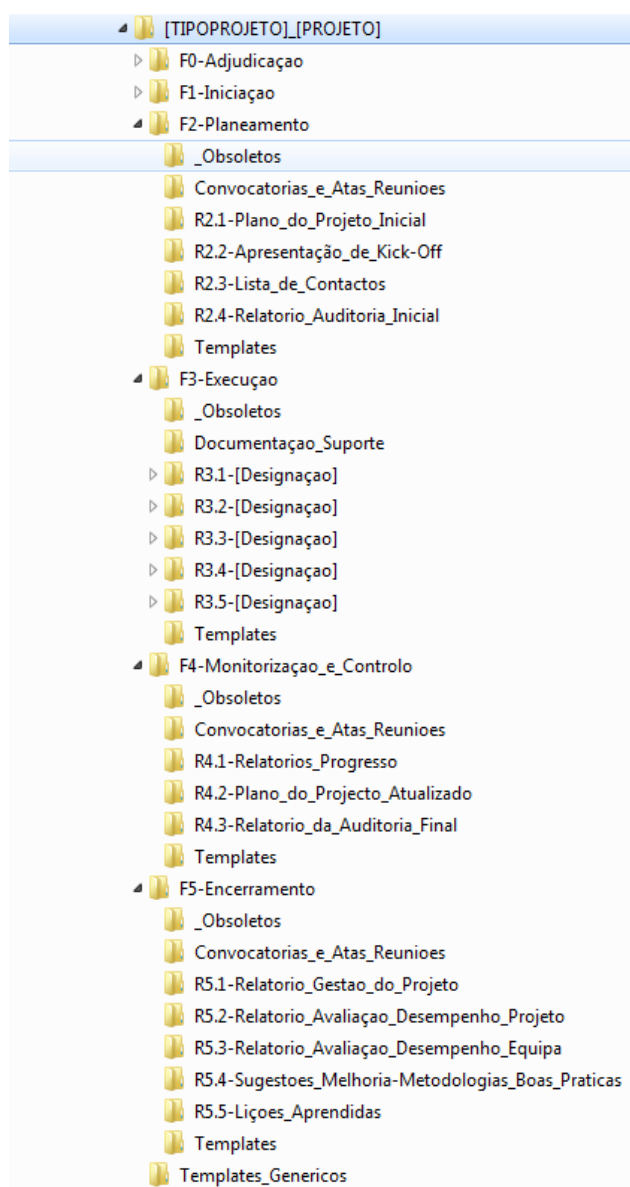


A organização por pacotes de atividades (as “work packages”) é o critério normal para a estruturação e organização da estrutura de classificação de cada projeto no INESC Porto, particularmente no sistema *Plone*. Durante as atividades do projeto, os conteúdos geridos na *Dropbox* são por vezes armazenados numa estrutura de pastas que não segue uma lógica definida, o que é obviamente prejudicial para a recuperação de informação. No entanto, o serviço de consultoria do INESC Porto tem-se esforçado em normalizar a estrutura hierárquica de pastas criada para gestão dos conteúdos de cada projeto, focando-se num critério que se aproxima do que é utilizado

Ilustração 6 - Estrutura de pastas em no Plone, mas numa perspetiva mais hierárquica utilizada no Plone

generalizada que se adequa a qualquer projeto – em vez de organizar as pastas de acordo com as atividades e tarefas do projeto, fá-lo de acordo com as fases típicas de gestão de um projeto (iniciação, planeamento, desenvolvimento, monitorização e controlo, e encerramento), o que permite uma estrutura idêntica em todos os projetos; as subpastas seriam depois organizadas de acordo com as atividades e tarefas específicas de cada projeto (ilust. 7)

Apesar de ser um critério um pouco diferente do que é utilizado no *Plone*, a sua base concetual é a mesma, ou seja, são ambos baseados na organização de atividades de um projeto, e orientadas para os resultados e tarefas necessárias para os alcançar. A utilização deste tipo de critérios é bastante lógica e intuitiva para o contexto de um projeto, uma vez que todos os envolvidos neste estão conscientes do planeamento das atividades do projeto e facilmente associam a estrutura hierárquica de pastas à organização do próprio projeto – isto permite perceber imediatamente qual é a classe – ou pasta – correspondente a cada conteúdo no momento do armazenamento deste no sistema, tornando fácil a utilização da classificação.



Mas será esta classificação útil para a gestão dos conteúdos neste

Ilustração 7 - Classificação hierárquica utilizada na Dropbox utilizada na área de consultoria do INESC Porto, adequada ao sistema através de uma estrutura hierárquica de pastas organizada segundo as fases de gestão de um projeto

contexto ou levantará problemas? Particularmente no Plone, e segundo os entrevistados, é eficaz.

“Normalmente quem está envolvido no projeto sabe exatamente em que parte da estrutura de pastas está o documento que precisa, [e chega até ele através de uma navegação pela estrutura]. Como cada colaborador sabe exatamente em que work package está exatamente a trabalhar, não se verificam problemas seja na submissão de um conteúdo ou na recuperação de informação.”

Anónimo 2

A estruturação da classificação segundo a própria organização das atividades do projeto parece ser eficaz para este tipo de contexto, e como é possível perceber pelo relato dos entrevistados, não origina problemas de recuperação de informação (no que diz respeito à recuperação dos conteúdos para gestão interna do projeto por parte de utilizadores envolvidos diretamente neste). No entanto, é necessário ter em conta que este contexto apresenta certas características específicas, e desse modo não é seguro afirmar que esse critério de divisão de classes seria eficaz noutro tipo de contexto com outras características. Além disso, temos de ter em conta que os entrevistados faziam parte dos colaboradores internos envolvidos diretamente no projeto – um colaborador externo ao projeto, ou mesmo um indivíduo externo à organização com permissões de acesso, poderia eventualmente sentir dificuldades em recuperar informação, uma vez que não conhecia a organização específica do projeto e, por conseguinte, a organização da estrutura de pastas. É importante pois definir quais as características em que este tipo de padrão de organização de conteúdos (classificação organizada de acordo com a própria organização do contexto) é eficaz. Então, poderemos dizer que uma classificação através de uma estrutura de pastas organizada de acordo com a própria organização do contexto seria eficaz num contexto com as seguintes características:

1. O contexto onde será implementada a classificação está já organizado por atividades, processos e tarefas independentes, e explicitamente delimitadas (as fronteiras entre atividades estão claramente definidas, e cada tarefa está enquadrada numa única atividade);
2. A quantidade de conteúdos a serem geridos não crescerá infinitamente até atingir quantidades imensuráveis (“Anónimo 1 - O facto do número de

documentos que criámos ser reduzido e a forma como a estrutura de pastas está organizada (atividades, resultados do projeto), faz com que a recuperação dos documentos seja simples e sem problemas”). Além disso, o tipo de conteúdos a armazenar no sistema de gestão de conteúdos estão pré-definidos;

3. A gestão dos conteúdos será apenas para utilização interna do projeto, e não para disponibilização a outros utilizadores que não os envolvidos diretamente no projeto.

O **aspecto 1** é bastante evidente, tendo em conta que é o critério utilizado para organizar a estrutura de pastas – o foco no estabelecimento de fronteiras claras entre atividades está relacionado com a necessidade de eliminação das ambiguidades e sobreposições entre atividades e tarefas, no sentido de não existirem dúvidas no momento de classificar cada conteúdo corretamente (no contexto deste tipo de projetos, “classificar o conteúdo corretamente” é entendido como o armazenamento do conteúdo na pasta respetiva à atividade/tarefa que deu origem a este). O **segundo aspecto** está relacionado com a inexistência de problemas de recuperação da informação testemunhada pelos entrevistados – o facto de não existirem problemas é compreensível se pensarmos que neste tipo de contexto a quantidade e tipo de objetos informacionais a serem criados estão pré-definidos, o que aliado à gestão dos conteúdos por classes derivadas diretamente das atividades do projeto, torna a recuperação da informação bastante intuitiva. A limitação da quantidade de conteúdos a serem armazenados é muito importante, uma vez que a recuperação típica neste tipo de sistema de classes é realizada através da navegação pela estrutura de pastas – se a quantidade de conteúdos crescesse infinitamente para números enormes, seria muito ineficiente recuperar informação através de pesquisa por navegação. Neste sentido, é importante incluir esta característica neste tipo de contexto, para garantir a eficácia e eficiência na recuperação de informação deste padrão de organização. O **terceiro aspecto** também está relacionado com a eficácia da recuperação, mas numa perspetiva diferente – a inexistência de problemas de recuperação de informação estará também relacionada com o facto de apenas ser considerada a recuperação da informação interna pelos próprios envolvidos no projeto, tal como afirmamos anteriormente. Como é lógico, uma estrutura hierárquica baseada nas atividades e tarefas do projeto não será intuitiva para outros utilizadores que desconheçam o contexto, e que muito

provavelmente não terão conhecimento da organização do projeto de forma a compreenderem a estrutura de pastas e a navegarem eficazmente pelas mesmas.

Neste sentido, é importante ter em conta que este padrão de organização de conteúdos apenas será útil e eficaz em contextos onde se pretenda a classificação da informação interna do projeto. Se existir a necessidade de recuperação eficaz de conteúdos por parte de utilizadores externos ao contexto (neste caso, utilizadores não envolvidos no projeto), é necessário recorrer a outro tipo de estrutura de classificação, de forma a garantir uma recuperação eficaz e eficiente por parte deste tipo de utilizadores.

Apesar de em alguns dos projetos, nos quais os entrevistados estiveram envolvidos, serem disponibilizados conteúdos ao público, estes não estavam classificados de acordo com um sistema de classes (e também não estavam armazenados na classificação em estrutura de pastas utilizada para a informação interna, uma vez que esta estava restrita aos colaboradores do projeto). Os conteúdos estavam, sim, organizados nas páginas do sistema, no entanto não foi atribuída qualquer classe a cada um destes conteúdos. Se a quantidade de conteúdos disponibilizada ao público for diminuta e estável ao longo do tempo, porventura não se verificarão grandes problemas de recuperação; no entanto é razoável pensar que num sistema onde haja necessidade de armazenar e atualizar grandes quantidades de conteúdo frequentemente, será bastante provável que a classificação dos conteúdos “públicos” passe a ser uma necessidade para evitar problemas graves de recuperação e de gestão da informação.

Tendo em conta este aspeto e a revisão teórica que fizemos acerca dos modelos estruturais de classificação, podemos perceber que a classificação facetada seria uma excelente solução para a classificação dos conteúdos “públicos”, num contexto onde este aspeto seja uma necessidade. **A classificação facetada parece ser uma excelente solução para a classificação deste tipo de conteúdos, uma vez que permite a maior flexibilidade e multidimensionalidade e, por conseguinte, será útil para mais utilizadores.** Além do mais, é um modelo estrutural bastante utilizado na gestão de conteúdos web (como em sites de vendas, entre outros) com qual os utilizadores já se encontram familiarizados, especialmente se o sistema de gestão de

conteúdos utilizado suportar uma interface de navegação pelas várias facetas. Se não existir esse suporte por parte do sistema, será necessário aplicar a classificação facetada de outra forma, como por exemplo através da atribuição de categorias aos conteúdos de acordo com uma classificação desenvolvida externamente – apesar de não ser o mais intuitivo, por vezes será necessário criar uma classificação externa ao sistema de forma a colmatar a falta de funcionalidades deste.

Relacionado com a questão da disponibilização de conteúdo ao público em geral, está também a questão cada vez mais importante da gestão do conhecimento, que neste caso significa, de forma simples, a partilha e transferência de conhecimento de projeto para outros relacionados – os colaboradores beneficiam com o conhecimento proveniente de um projeto anterior, no entanto este não é muitas vezes gerido de forma adequada para permitir a sua recuperação por utilizadores que não estiveram diretamente envolvidos no projeto.

“Se estivermos a falar da organização dos conteúdos no decurso do projeto, acho que essa forma de organizar é a que se adequa melhor à gestão de processos do projeto, que estão organizadas em work packages. Agora, se estivermos a pensar em utilizar essa informação pós-projeto, penso que já não é adequado.”

Anónimo 3

E é este um dos principais problemas de gestão de informação no INESC Porto segundo o serviço de consultoria – apesar de não existirem grandes problemas na recuperação de conteúdos de um projeto por parte de utilizadores diretamente envolvidos neste, o caso muda de figura quando estamos a falar de recuperação de conteúdos por parte de utilizadores que não estão diretamente envolvidos no projeto. Mesmo que a estrutura hierárquica de pastas de determinado projeto na Dropbox esteja organizada segundo as boas práticas indicadas pelo serviço de consultoria (o que não acontece sempre, uma vez que os colaboradores do projeto não seguem sempre estas boas práticas), e tenham, assim, uma estrutura familiar para um utilizador conhecedor do ambiente típico de um projeto (organização por fases do processo de um projeto), surgem várias vezes necessidades informacionais que não são satisfeitas de forma eficaz e eficiente através da típica navegação por pastas. Isto é, se um utilizador estiver, por exemplo, interessado em recuperar informação sobre projetos na área do

“calçado” no universo geral de projetos do INESC Porto, terá a missão bastante dificultada porque os conteúdos estão classificados de acordo com as atividades do próprio projeto, e carecem de perspetivas mais gerais e diversificadas. Os utilizadores teriam, assim, de recorrer à pesquisa por navegação, o que se revela extremamente moroso quando estamos a falar de um sistema com uma quantidade de pastas muito elevada. Isto revela-se um grande obstáculo à recuperação eficaz de informação, e consequentemente, à própria gestão do conhecimento na organização. Esta é, então, uma das maiores preocupações da gestão de informação no INESC Porto, e para o qual a classificação poderá ser uma mais-valia. **Neste caso, a classificação facetada dos conteúdos de cada projeto seria uma solução bastante útil, permitindo classificar os conteúdos de acordo com várias perspetiva, e potenciando dessa forma a recuperação dos conteúdos por parte de utilizadores externos ao projeto em questão, e também melhorando a partilha de conhecimento dentro da organização.** A eficácia da recuperação dependeria, obviamente, da adequação das facetas às necessidades informacionais destes utilizadores externos.

Posto isto, podemos dizer que as diferentes necessidades de recuperação de informação que podem existir nos utilizadores de determinado contexto organizacional podem ser relacionadas com padrões de organização e classificação de conteúdos adequados à sua satisfação.

Necessidades de recuperação de informação	Padrão de organização e classificação de conteúdos adequado
<i>Necessidade de recuperação de informação para gestão interna do contexto organizacional, de acordo com a estrutura do contexto</i>	O modelo estrutural em hierarquia (de pastas), de acordo com as atividades, processos e/ou tarefas realizadas no contexto é o mais adequado. É importante ter em conta que se a quantidade de conteúdos a gerir no sistema é relevante, uma vez que poderá tornar a pesquisa por navegação bastante ineficiente. Nestes casos poderá ser necessário um modelo estrutural mais complexo, como a classificação facetada, que permite a recuperação por introdução de interrogações de pesquisa.

<i>Necessidade de recuperação de informação a disponibilizar ao público, por utilizadores externos ao contexto</i>	O modelo estrutural de classificação facetada é adequado, onde a utilização de facetas permite corresponder às possíveis perspetivas do utilizador externo ao contexto com necessidade de recuperação dos conteúdos “públicos” disponibilizados pelo contexto.
<i>Necessidade de recuperação de informação no âmbito da gestão do conhecimento para o futuro –</i>	O modelo estrutural de classificação facetada é adequado, com a utilização das facetas necessárias para corresponder às possíveis necessidades do utilizador-colaborador do contexto.
<i>Necessidade de recuperação de informação de acordo com necessidades pessoais e de manutenção de um ambiente incentivador da colaboração e participação na partilha de conhecimento</i>	O modelo estrutural da folksonomia é adequado, permitindo que o utilizador escolha e atribua tags ao conteúdo de acordo com as suas perspetivas pessoais, contribuindo ativamente com a sua individualidade para a gestão do conhecimento

Esta relação necessidade de recuperação – padrão de organização permite que cada contexto organizacional possa perceber de que forma pode adaptar a estrutura de uma classificação às necessidades específicas dos seus colaboradores.

3.2 Adaptação da classificação ao sistema de gestão de conteúdos

Os sistemas utilizados na gestão dos conteúdos representam um importante fator na futura eficiência e eficácia de um sistema de classificação – são as funcionalidades das plataformas que permitem que a classificação possa ser aplicada e utilizada. As

funcionalidades do sistema de gestão de conteúdo têm então um impacto direto na utilidade e usabilidade de um sistema de classificação e, desse modo, devem ser feitos esforços na tentativa de adaptar, da melhor forma possível, o modelo estrutural de classificação às funcionalidades do sistema de gestão de conteúdos que suportará a utilização da classificação.

Como já foi dito no início deste capítulo, no contexto do INESC Porto utilizam-se principalmente duas plataformas, o *Plone* e a *Dropbox*, segundo o que foi possível apurar nas entrevistas. No que diz respeito ao *Plone*, esta plataforma possui, além da

função *standard* de criação e estruturação hierárquica de um conjunto de pastas, uma opção de atribuição de categorias a cada conteúdo/pasta – categorias que devem ser criadas pelos utilizadores que possuam permissões para tal (ilust. 8). No momento de atribuição de uma categoria a um conteúdo, o utilizador pode criar uma nova categoria nesse

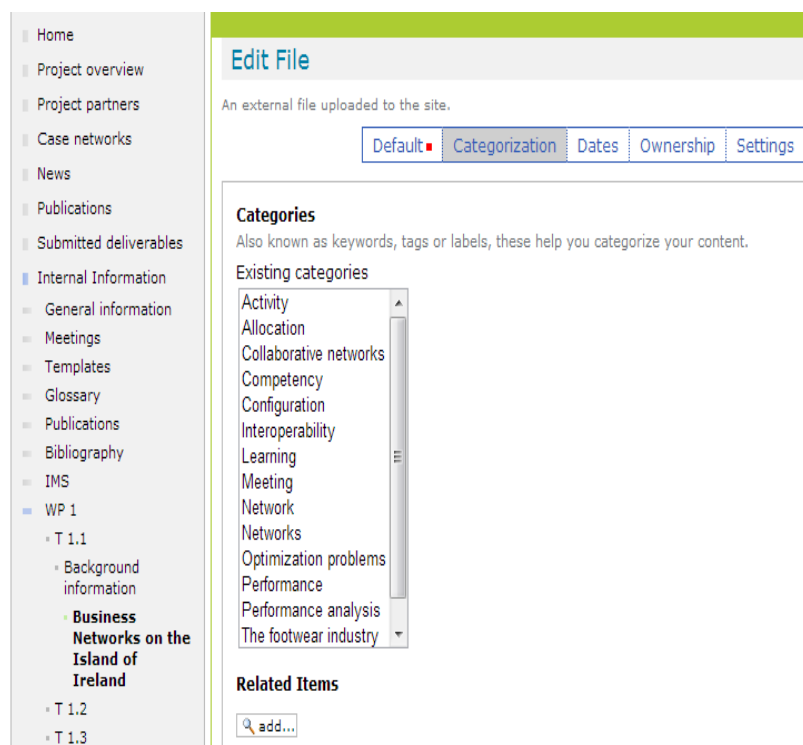


Ilustração 8 - Funcionalidade de atribuição de categorias do Plone, um dos possíveis métodos para "classificar" o conteúdo da plataforma

momento, ou utilizar uma categoria já existente (que possivelmente já havia sido utilizada para a categorização de outros conteúdos). Como foi possível perceber durante a análise dos padrões de estruturação do sistema de classes, as funcionalidades de categorização das plataformas não são normalmente utilizadas pelos colaboradores dos projetos; a classificação e organização de conteúdos apenas “aproveitam” as opções de criação e organização de pastas, e de armazenamento dos conteúdos nestas.

No que diz respeito à Dropbox, a situação é muito semelhante: a Dropbox é uma aplicação que funciona no ambiente do Windows Explorer, e este último também permite a criação e organização de pastas, e possui um sistema de atribuição de etiquetas a cada pasta/conteúdo (ilust. 9).

As classificações utilizadas normalmente nos projetos do INESC Porto estão normalmente estruturadas segundo um modelo hierárquico, o que permite uma adaptação direta e bastante simples à estrutura de pastas – existe apenas uma transição de

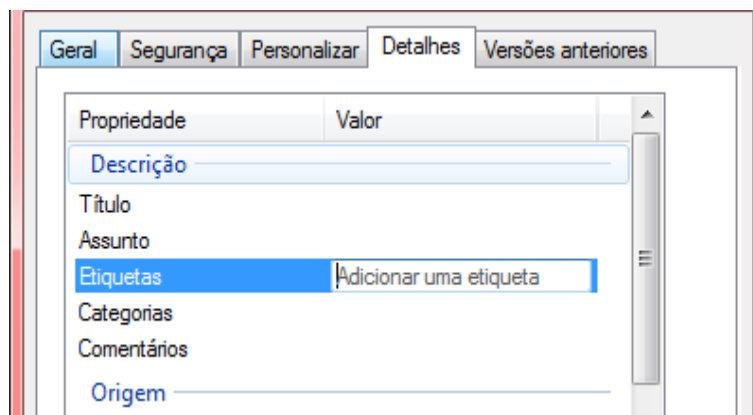


Ilustração 9 - Funcionalidade de atribuição de etiquetas a conteúdos do Windows Explorer

uma hierarquia de classes para uma hierarquia de pastas. Neste prisma, as funcionalidades de criação e organização de pastas parecem ser uma opção intuitiva e eficaz.

Os problemas surgem quando é necessário adaptar uma classificação com um modelo estrutural mais complexo, como por exemplo o modelo por facetas. Uma classificação facetada não permite uma adequação direta e intuitiva a uma estrutura de pastas – seria necessário organizar uma hierarquia de pastas para cada faceta, e no momento da classificação, o utilizador teria de criar várias cópias do mesmo conteúdo de forma a poder “classificá-lo” de acordo com cada faceta (ou seja, teria de armazenar cada cópia do conteúdo em cada uma das hierarquias de pastas, correspondentes a cada uma das facetas). Apesar dos contextos em que os entrevistados estiveram envolvidos não terem feito uso de classificações facetadas, é seguro afirmar que um sistema por facetas tornaria muito mais complexa, não só a adaptação da classificação ao sistema, mas também a própria classificação do conteúdo pelo utilizador. E é aqui que as funcionalidades de atribuição de categorias/etiquetas das duas plataformas podem ser bastante úteis, permitindo com que o utilizador possa atribuir categorias ao conteúdo de acordo com o conjunto de facetas da classificação. Com um esforço mínimo, o utilizador pode classificar o conteúdo em várias perspetivas, através da

atribuição de várias categorias ao conteúdo, cada uma delas correspondente a uma das facetas da classificação criadas para o contexto. As categorias/etiquetas, quando complementadas por uma estrutura hierárquica de pastas (onde esta poderá representar a faceta mais importante para o contexto), permite que os utilizadores possam fazer pesquisas através da navegação e, também, através da introdução de interrogações de pesquisa no motor de busca do sistema de conteúdos. Isto traduz-se numa melhor usabilidade do sistema, e, principalmente, numa provável melhoria da recuperação da informação.

É necessário ter em conta que estas duas plataformas ou aplicações *não* são os exemplos com maior quantidade de funcionalidades. Outros sistemas de gestão de conteúdos mais complexos apresentam muitas mais opções em termos de customização e classificação de conteúdos como o *Microsoft Sharepoint*. As wikis semânticas são outro tipo de plataforma que permite a adaptação de uma classificação ao sistema de organização do conhecimento da wiki, de forma a organizar as páginas da wiki segundo a classificação desenvolvida. No entanto, nem todos os contextos organizacionais têm a possibilidade de implementar os sistemas mais adequados e, desse modo, terão de fazer mais esforços no momento da adaptação das classificações aos sistemas tecnológicos.

Posto isto, é importante ter em conta que a adaptação da classificação à plataforma tem uma influência tremenda na usabilidade e utilidade da classificação. Neste sentido, não só é importante desenvolver uma classificação adequada ao contexto, mas também descobrir como adaptá-la da melhor forma possível ao sistema de gestão de conteúdos. A seleção da tecnologia não é, muitas vezes, uma opção num processo de desenvolvimento de classificação (já está implementada e não pode ser alterada), no entanto, quando isso for possível, devem ser consideradas e analisadas as opções mais adequadas em termos de plataformas para o tipo de classificação que se pretende implementar. Então, podemos dizer que a utilização da funcionalidade de atribuição de categorias/etiquetas parece ser a opção mais lógica e eficaz na adaptação dos modelos estruturais de classificação mais complexos aos sistemas de gestão de conteúdos, contudo deverão ser consideradas outras e melhores possibilidades quando os contextos ou as plataformas assim o exigem e/ou permitem.

3.3 Definição dos conceitos, termos e relações

Como vimos na secção “estruturação do sistema de classes”, o facto do modelo estrutural de classificação utilizado na maioria dos casos ser uma hierarquia baseada nas próprias (grupos de) atividades do projeto, fez com que os envolvidos no projeto não necessitassem obrigatoriamente de ponderar sobre qual o modelo estrutural de classificação a implementar no sistema de gestão de conteúdos. O mesmo acontece com a definição dos conceitos e termos a representar na classificação: como a estrutura típica dos contextos (isto é, os projetos onde INESC Porto está envolvido) está já organizado de uma forma hierárquica e lógica (seja através do conceito de Work Packages, ou através do modelo hierárquico de fases do projeto desenvolvido pelo serviço de consultoria), a definição dos conceitos e termos a incluir na classificação acaba por se transformar numa mera transposição da organização das atividades do contexto para uma estrutura hierárquica de classificação (onde, como vimos, cada classe representa uma atividade/processo/tarefa real do contexto).

“Nós temos um procedimento que, ao iniciar um novo projeto, são criados automaticamente a estrutura de pastas de alto nível e os templates para os documentos. Isso já é do conhecimento de todas as pessoas que estão a trabalhar no projeto, e portanto não existe discussão quanto a esses temas. Houve no passado, mas neste momento já toda a gente sabe que a estrutura está estabilizada.”

Anónimo 1

Como a maioria dos contextos acaba por ter estruturas bastante similares, este “processo” de definição de conceitos e termos torna-se numa tarefa quase automatizada (assim com a definição do modelo estrutural de classificação). Apesar de este processo não levantar problemas de recuperação de informação por parte dos utilizadores diretamente envolvidos no projeto, é importante mais uma vez ter em conta que o mesmo não se passa com os utilizadores externos ao projeto. Estes últimos necessitam que os conteúdos sejam classificados sob diferentes perspetivas para além daquelas implicitamente definidas pelo contexto interno do projeto.

Posto isto, é importante ter em conta que, mais uma vez, os contextos com necessidades de recuperação mais avançadas fazem com que a definição dos conceitos e termos a representar na classificação seja uma tarefa mais complexa que aquela que foi possível constatar pelos relatos dos entrevistados. A complexidade da definição dos

conceitos e termos é, então, influenciada pela complexidade do modelo estrutural da classificação, e ambas estão dependentes das características do contexto no qual a classificação será implementada.

Então, e pegando nas ideias que alcançamos na análise dos padrões de estruturação do sistema de classes, podemos dizer que a definição de conceitos e termos a representar num modelo estrutural em hierarquia (organizada de acordo com as atividades/fases do projeto) será uma tarefa bastante simples e imediata, tal como foi relatado pelos entrevistados. Contudo, a complexidade surgirá quando for necessário definir os conceitos e termos para modelos estruturais mais complexos, como a classificação facetada. Neste modelo estrutural, as várias facetas requerem que os colaboradores, encarregues da tarefa de definir os conceitos e termos, “olhem” para o contexto e para os conteúdos a gerir através de várias perspetivas e caracterizem os mesmos de acordo com as facetas que pré-definiram para a classificação. A complexidade da tarefa variará assim consoante diversos fatores, entre os quais a natureza e quantidade das facetas, a homogeneidade/heterogeneidade e assunto dos conteúdos, e como já foi mencionado, as necessidades de recuperação de informação do contexto.

As facetas poderão optar por perspetivas relacionadas com:

- ✓ Características do contexto (p. ex. *atividade, processo, tarefa*);
- ✓ Características “físicas” dos objetos informacionais (p. ex. *tipo* (de documento ou conteúdo), *versão*);
- ✓ Características do assunto que trata os conteúdos, assim como outras que se revelem como relevantes para os utilizadores do sistema.

O tipo de facetas a incluir nas classificações é, por isso, difícil de normalizar e problematizar, devido às diversas variáveis que influenciam as perspetivas segundo as quais podemos analisar o conteúdo. Mesmo as tentativas de definição de “facetas fundamentais” de Ranganathan e Vickery (citados em Tennis, 2011 e Broughton, 2011) não se adequam a contextos profissionais práticos, onde as competências dos profissionais não incluirão, na maioria dos casos, habilitações específicas e técnicas de classificação de informação, e onde o típico utilizador comum não compreenderia imediatamente o que significariam essas “facetas fundamentais”, uma vez que a sua

designação não estaria sempre diretamente relacionada com o seu significado num dado contexto específico. Neste sentido, a definição das facetas requer alguma capacidade de análise as características do contexto e do conteúdo e de seleção daquelas que se mostrem como possivelmente relevantes para o futuro utilizador do sistema. Para além de senso comum e pensamento lógico, claro.

Apesar da dificuldade em normalizar a definição de facetas e os termos a utilizar para descrever o conteúdo de acordo com cada uma destas, existem alguns pontos que podem auxiliar o processo:

- ✓ Criar facetas sempre de acordo com as possíveis perspetivas que o utilizador comum considerará relevante para as suas necessidades de informação;
- ✓ As facetas relacionadas com as características do contexto e as características físicas dos objetos informacionais são normalmente mais relevantes para os utilizadores do contexto interno (estejam eles diretamente ligados ao projeto ou façam apenas parte da mesma organização); as facetas relacionadas com o assunto dos conteúdos serão mais importantes para os utilizadores externos ao contexto, uma vez que não estão familiarizados com as características deste;
- ✓ O número de termos utilizados em cada uma das facetas não se deve tornar demasiado elevado, uma vez que a elevada especificidade torna as fronteiras entre termos/conceitos demasiado turvas, o que por sua vez aumenta a complexidade do processo de classificação do conteúdo. Para evitar a elevada especificidade, devem ser criados novos termos que substituam conjuntos de termos cujos conceitos que representam fossem próximos.
- ✓ A consulta de classificações já disponíveis (como tesauros) poderão ser uma excelente opção para se definirem os termos, ou mesmo facetas, relacionados com o assunto dos conteúdos. Algumas normas de criação de tesauros (como a NP 4036¹⁰, ISO 25964¹¹) também poderão ser muito úteis para ajudar no processo de criação de termos.

¹⁰ PORTUGAL. Instituto Português da Qualidade. Comissão Técnica 7 (1993). Norma Portuguesa 4036 : documentação : tesauros monolíngues : diretivas para a sua construção e desenvolvimento. Lisboa : I. P. Q.

¹¹ INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (2011). Information and documentation -- Thesauri and interoperability with other vocabularies -- Part 1: Thesauri for information retrieval. França: ISO.

Por último, é importante também referir que a indexação automática poderá ser utilizada como alternativa à classificação dos assuntos dos conteúdos, pelo menos quando a necessidade de recuperação de informação por assunto não for uma prioridade. Isto acontecerá, normalmente, em contextos onde não existirem conteúdos públicos ou onde a gestão de conhecimento para o futuro não for relevante – nestes casos, a classificação dos conteúdos segundo as características físicas dos objetos informacionais e do contexto será, em princípio, suficiente para a recuperação de informação por parte dos utilizadores. Contudo, e mais uma vez, estas decisões deverão ser ponderadas tendo em conta as características específicas de cada contexto.

Concluindo, a definição dos conceitos e termos a representar na classificação é uma tarefa também dependente das necessidades de recuperação de informação, no entanto a diversidade no que diz toca às características do contexto, às características físicas dos objetos informacionais e à natureza do assunto dos conteúdos, faz com que seja bastante complicado, e até irrealista, definir um processo normalizado com soluções apropriadas para padrões definidos. Deste modo, a melhor opção é optar por uma série de recomendações práticas para auxiliar o processo, mas onde se ressalva a importância da capacidade de análise crítica do contexto e do conteúdo e de senso comum por parte dos desenvolvedores da classificação.

4. Processos colaborativos para o desenvolvimento e gestão de classificações

O típico contexto organizacional é definido não só pela natureza da sua área de atividade, mas também pelas características dos colaboradores, pela tecnologia que utiliza e pelos processos organizacionais - a interação entre estes três elementos influencia a forma como a gestão de informação é realizada. A interação existente neste trinómio não acontece, no entanto, apenas entre os três elementos: a complexidade inerente a cada um dos elementos – pessoas, tecnologia e processos – exige que existam inter-relações entre os seus subelementos. Ou seja, e dando um exemplo, não só é necessário os processos integrarem a tecnologia e a atividade humana, mas também é fulcral que os vários processos estejam interligados de forma a garantirem eficiência e eficácia no contexto organizacional. A mesma necessidade existe com a tecnologia e a sua interoperabilidade.

O elemento humano – o conjunto dos colaboradores do contexto, e num nível mais abrangente, todos os colaboradores - é, porventura, aquele cuja interação entre subelementos (cada colaborador individual) é mais complexa e difícil de controlar. A individualidade mental inerente a cada ser humano impede que se possa prever e controlar eficazmente a forma como os colaboradores devem interagir durante os processos organizacionais. A organização pode influenciar o comportamento colaborativo através de uma cultura adequada, no entanto essa influência não será eficaz em todos os colaboradores e estará, inevitavelmente, dependente das características individuais de cada um destes. A colaboração no contexto organizacional assume assim vital importância não só na interação entre colaboradores, mas também na própria inter-relação com os restantes dois elementos do trinómio, nomeadamente tecnologia e processos. Os processos colaborativos são assim muito importantes para a gestão da informação: a sua integração nos sistemas tecnológicos e o seu papel vital nos processos organizacionais só é possível através da atividade humana, que desenvolve e controla a forma como estes três elementos se relacionam, procurando sempre melhorar a eficiência e a eficácia de acordo com os objetivos do contexto organizacional.

Por conseguinte, a classificação, como aspeto essencial no ciclo de gestão da informação, beneficia com a existência de um ambiente de colaboração. Neste sentido, é importante que se definam processos colaborativos que suportem o desenvolvimento e gestão de classificações, para que o método, e a sua futura aplicação, sejam capazes de desencadear processos colaborativos não só no desenvolvimento e gestão da classificação, mas também, no melhor cenário, na própria gestão da informação.

Tal como no último capítulo, também neste vamos analisar a informação recolhida nas entrevistas aos colaboradores do INESC Porto, assim como aquela que foi possível retirar durante a revisão de literatura, de modo a ser possível chegar a uma série de conclusões acerca da melhor forma de influenciar a atividade colaborativa no processo de desenvolvimento e gestão de classificações. Estas conclusões serão extremamente importantes para o desenvolvimento do método.

4.1 Desenvolvimento colaborativo da classificação

No último capítulo, verificou-se que na maioria dos projetos onde os entrevistados estiveram envolvidos, tanto a seleção do modelo estrutural de classificação, como a definição dos conceitos e termos, foram levadas a cabo através de um processo quase “automatizado”, que não envolveu grande discussão por parte dos colaboradores do projeto (no que diz respeito à gestão de conteúdos).

Tendo em conta este aspeto, foi possível deduzir, durante as próprias entrevistas que, na grande maioria dos projetos, a construção das estruturas hierárquicas de pastas para a gestão de conteúdos não foi fruto de processos colaborativos entre os envolvidos no projeto -como verificamos no capítulo anterior, o processo de desenvolvimento da classificação passou pela simples transposição da organização do projeto para uma estrutura de pastas, e foi levado a cabo normalmente pelo responsável para instalação dos sistemas de gestão de conteúdos (pelo menos no que diz respeito a gestão de conteúdos na plataforma *Plone*). No que diz respeito à classificação de conteúdos na *Dropbox*, o trabalho do serviço de consultoria do INESC Porto na tentativa de normalizar as estruturas hierárquicas de pastas utilizadas em cada projeto ajuda também a que a classificação nessa plataforma seja cada vez mais vezes feita de forma “automatizada”, e sem necessidade de discussão ou trabalho colaborativo. Aliás, o

próprio serviço de consultoria está encarregue de corrigir possíveis erros cometidos, pelos colaboradores de projetos, na estruturação das pastas e organização dos conteúdos. Estas correções trazem benefícios para a organização geral da estrutura de pastas geral, no entanto retira também alguma das responsabilidades dos contextos específicos em desenvolverem adequadamente as suas classificações – a situação ideal para um contexto como o do INESC Porto (onde temos contextos específicos – cada projeto -, dentro de um contexto maior – INESC) seria uma em que os colaboradores de cada projeto seguissem regras gerais de classificação definidas pelo serviço de consultoria, para que a organização e classificação dos conteúdos fosse normalizada ao ponto de se tornar intuitiva e natural para os colaboradores; isto traduzir-se-ia por sua vez numa plataforma *Dropbox* melhor organizada, onde todos os projetos possuíam uma organização o mais homogênea quanto possível, o que, consequentemente, resultaria numa melhor recuperação de informação por parte dos utilizadores, pelo menos no que diz respeito à recuperação de informação através de navegação por pastas. Em contextos com dimensões mais reduzidas e áreas de atividade mais específicas é lógico que a complexidade dos aspetos relacionados com a classificação não seja tão elevada, no entanto a colaboração e a normalização no processo de classificação é sempre vital para a utilidade do sistema de classes e para a eficácia e eficiência da recuperação da informação.

A motivação para o trabalho colaborativo é um aspeto que está sempre dependente do complexo “fator humano”, e desse modo é difícil de prever de que forma as decisões e atividades promovidas serão benéficas para a coesão e colaboração entre os colaboradores. O mesmo acontece numa equipa com a responsabilidade de desenvolver uma classificação – é impossível prever a eficácia das iniciativas promovidas na melhoria da colaboração entre elementos da equipa. Esta situação piorará, logicamente, se a equipa não possuir competências adequadas para realizar as atividades que lhes são pedidas; é óbvio que nem todos os contextos serão escassos em termos de competências de gestão da informação, no entanto o método que se pretende desenvolver neste trabalho pretende privilegiar a agilidade de aplicação a qualquer contexto, e, desse modo, serão mais úteis atividades simples que se possam realizar em qualquer contexto, independentemente da qualidade e quantidade de competências existentes. A motivação para colaboração num processo de desenvolvimento de uma classificação terá, assim, de passar pela realização de atividades colaborativas com uma

complexidade adequada às características dos colaboradores, para que cada colaborador se possa sentir importante no atingir dos objetivos, e para que seja dada mais relevância às decisões individuais de cada elemento da equipa.

Uma das soluções mais interessantes para motivar o trabalho colaborativo neste tipo de processo é a construção de uma rede semântica de conceitos inter-relacionados e relevantes para o contexto. A construção de um artefacto simples como um mapa de conceitos permite que cada elemento possa contribuir com o seu conhecimento e perspectivas individuais acerca do contexto, podendo representar estas através de um desenho de compreensão e realização relativamente fácil.

“ [...] permite que duas equipas trabalhem primeiro individualmente nas suas ideias e no seu mapa de conceitos, e depois a ferramenta permite agregar os dois mapas. O mapa de conceitos pode depois ser utilizado para criar uma classificação para gerir um sistema de gestão de conteúdos. O conceptME não tenta apenas ser uma ferramenta de automatização, mas sim algo que incentive à discussão. Isto é importante porque mesmo dois especialistas da área podem ter duas visões diferentes do conhecimento do domínio, e assim a ferramenta pretende também apoiar essa discussão.”

Anónimo 3

Os mapas de conceitos permitem que cada colaborador possa exprimir as suas decisões e conhecimento de forma simples, e que possa sentir o valor das suas decisões na construção da rede semântica final. Além disso, existem vários *softwares* que auxiliam esta tarefa, como por exemplo o *CmapTools* e o *ConceptME*: este último, particularmente, permite a combinação de vários mapas de conceitos para a formação de um só que abarque todas as perspectivas individuais, eliminando as redundâncias simultaneamente. Este tipo de *software* permite tornar a construção deste tipo de mapas numa tarefa mais agradável, através de uma interface simples - mais uma vez, não só é importante motivar a colaboração através de atividades que mostrem a importância do colaborador para atingir os objetivos, mas também tornar essas atividades o menos morosas e complexas possível.

O resultado da combinação dos vários mapas de conceitos é interessante não só no ponto de vista colaborativo, mas também na medida em que permite conhecer as várias perspectivas nas quais o contexto pode ser compreendido e representado. Neste

sentido, é importante que os colaboradores possam ter liberdade para “olharem” para o contexto de acordo com as perspetivas que entenderem, e o representem de acordo com essa visão. Uma rede semântica rica em perspetivas permite que a mesma possa ser decomposta em várias “redes” mais específicas, que podem depois servir de base para o desenvolvimento das classificações que satisfaçam as necessidades de recuperação de informação.

Este tipo de atividade será tão eficaz quanto maior for a discussão dos vários conceitos entre os elementos da equipa. É importante que a equipa de desenvolvimento da classificação procure discutir não só a estrutura do sistema de classes, mas também os próprios conceitos e relações a representar na mesma. Para isto são necessárias plataformas colaborativas, como fóruns ou mesmo wikis semânticas, para que todos os colaboradores possam partilhar as suas ideias no mesmo espaço virtual, de forma a não só incentivar a discussão, mas também a registar todo o conhecimento produzido.

É importante referir nem todos os contextos terão os recursos humanos e temporários para “gastar” numa tarefa como a de construção de uma rede semântica. Particularmente em contextos mais complexos, com várias dezenas de colaboradores envolvidos em diferentes “subcontextos”, será porventura mais acessível e realista direccionar os esforços principais de desenvolvimento de uma classificação para os departamentos responsáveis pela gestão dos conteúdos do contexto. Por exemplo, no contexto complexo do INESC Porto, o serviço de consultoria tem maiores responsabilidades na gestão dos conteúdos e muitas das melhorias na recuperação de informação partirão regularmente das suas iniciativas. Neste prisma, a importância da colaboração entre os vários colaboradores do contexto deixa de ser tão elevada no momento do desenvolvimento da classificação, mas passa a ser essencial na sua posterior utilização e gestão.

4.2 Utilização e gestão colaborativa da classificação

Um desenvolvimento colaborativo bem-sucedido de uma classificação é inútil se não se formarem bases para uma posterior utilização colaborativa da mesma, entre todos os colaboradores do contexto. É por isso importante que a equipa envolvida no desenvolvimento da classificação pondere sobre possíveis formas de motivar a

utilização adequada desta, bem como de normalizar este processo, sempre com o intuito de tornar o ato de “classificar conteúdo” numa tarefa intuitiva e natural, e não numa tarefa morosa que perturbe o colaborador.

Num contexto com uma quantidade de colaboradores reduzida, é possível que todos os colaboradores estejam envolvidos na fase de desenvolvimento da classificação, e, como tal, não existiriam grandes dificuldades em motivar uma utilização colaborativa da mesma – o envolvimento no processo de desenvolvimento permite que os colaboradores percebam a importância do que estão a criar, os seus benefícios e de forma pode ser utilizada, o que se traduzirá, num ponto de vista lógico, numa maior motivação para o seu uso. No entanto, mesmo neste tipo de contextos, podem existir mudanças a nível dos colaboradores, e é importante que a equipa envolvida no desenvolvimento da classificação garanta que os possíveis novos colaboradores possam compreender e usar a classificação de forma adequada. Num contexto de grandes dimensões, onde o processo de desenvolvimento de uma classificação seria normalmente realizado pelo departamento com maiores responsabilidades na gestão de conteúdos (dando, mais uma vez, o exemplo do INESC Porto e do seu serviço de consultoria), é ainda mais importante que sejam tomadas medidas para motivar e normalizar a utilização colaborativa e adequada da classificação, uma vez que muitos dos colaboradores não estiveram envolvidos no processo de desenvolvimento da mesma.

“É importante que haja uma divisão do esforço, entre tarefas de indexação que podem ser automatizadas e tarefas de classificação por parte dos colaboradores, e que haja um investimento na criação de ajudas à classificação que facilitem o trabalho o mais possível às pessoas, que não as perturbem mesmo. É importante que se motive uma cultura de utilização da classificação. Isto também está relacionado com a própria funcionalidade e usabilidade dos sistemas, claro.”

Anónimo 3

A criação de um manual de utilização da classificação é possivelmente uma das soluções mais simples e eficazes para normalizar este processo, permitindo que os colaboradores pudessem facilmente perceber quais os passos que precisavam de realizar no momento de armazenamento de conteúdos no sistema. O serviço de

consultoria do INESC Porto utilizar uma solução semelhante para normalizar os nomes dos documentos criados no âmbito de cada projeto, e segundo o que foi possível perceber, parece ser realmente uma solução viável para a normalização deste tipo de tarefas. No que diz respeito à motivação para a utilização da classificação, é uma questão mais complexa que envolve vários fatores, entre os quais a competência da liderança, o tipo de cultura do contexto e as características individuais dos colaboradores. No entanto, valoriza-se, obviamente, a comunicação entre diretores-colaboradores na tentativa de introduzir o processo de classificação nas rotinas de trabalho.

Outro aspeto muito importante está relacionado com a gestão e atualização da classificação. É importante que a equipa que esteve envolvida no processo de desenvolvimento fique atenta ao feedback dos utilizadores da classificação, de forma a poder monitorizar possíveis novas necessidades de recuperação de informação, incapacidades da classificação (termos insuficientes para classificar adequadamente um conteúdo) ou mesmo erros na semântica do sistema de classes (termos e inter-relações mal enquadrados). Este feedback será essencial para a equipa poder proceder às atualizações necessárias, sejam elas possíveis adições de novos termos ou, mesmo, a criação/alteração do modelo estrutural de classificação para corresponder a uma nova necessidade de recuperação de informação. Além disso, é importante que a equipa esteja atenta a possíveis erros na atribuição da classificação. Esta análise de feedback e consequentes atualizações estabeleceriam uma espécie de “gestão da qualidade” da classificação, que seria essencial para manter níveis de recuperação de informação adequados às necessidades do contexto.

Este tipo de feedback e gestão colaborativa da classificação requer, contudo, meios de comunicação adequados para que cada colaborador possa fazer a sua contribuição. Uma boa solução seria a utilização de uma plataforma online colaborativa, como uma *wiki* ou o fórum dos próprios sistemas de gestão de conteúdos, onde o feedback pudesse ser registado e consultado pelos outros colaboradores e pela equipa responsável pelas atualizações da classificação, e onde também fosse possível a discussão do mesmo. Esta plataforma colaborativa seria, idealmente, já utilizada pela equipa para comunicar durante o processo de desenvolvimento da classificação, sendo depois transposta para funções de plataforma colaborativa para a gestão da qualidade

da classificação. É, por isso, fulcral que exista uma fase de planeamento anterior ao próprio desenvolvimento da classificação, com o objetivo de definir não só a equipa responsável, mas também quais as plataformas que serão utilizadas para suportar o trabalho colaborativo e a comunicação. É importante também que os conteúdos criados nas plataformas colaborativos pelos colaboradores estejam organizados de acordo com uma classificação, se possível semelhante àquela desenvolvida para o sistema de gestão de conteúdos. Desta forma, os novos utilizadores, que desconheçam o contexto e o seu domínio, poderão consultar as plataformas colaborativos onde estará registado o conhecimento proveniente das discussões. Este tipo de aprendizagem, aliado à consulta da classificação, permite uma gestão do conhecimento mais preocupada com a usabilidade do utilizador e com as suas necessidades de recuperação de informação.

“É lógico que um utilizador que chega ao sistema de gestão de conteúdos de um projeto não irá perceber o domínio do conhecimento através das work packages. Se ele pudesse aprender com a taxonomia desse sistema, ou mesmo com a informação registada no fórum, seria muito melhor.”

Anónimo 3

4.3 Planeamento do processo de desenvolvimento e gestão da classificação

A importância do planeamento para a gestão de qualquer processo com algum grau de complexidade é evidente – um processo é normalmente composto por várias fases e etapas, cada uma com metas, recursos e resultados esperados específicos, e todos estes elementos devem ser tidos em conta antes do início do projeto de forma a se poder garantir o processamento correto das tarefas e atividades. O desenvolvimento e gestão de uma classificação é também um processo que, apesar de não ter a mesma complexidade de um processo industrial, inclui igualmente várias pequenas tarefas diferentes para se atingir o objetivo principal – é necessário, assim, que todos estes elementos sejam planeados e organizados de forma a se criarem bases para o sucesso do processo.

Em muitos aspetos, a fase de planeamento do processo de desenvolvimento de uma classificação inclui as tarefas existentes num “planeamento” usual, nomeadamente o estabelecimento dos objetivos e benefícios do resultado do processo (a classificação) e a calendarização das tarefas. Além disso, a definição dos recursos também é uma das

tarefas mais importantes no planeamento deste processo: não só é necessário definir a equipa encarregue de desenvolver e gerir a classificação, mas também selecionar (e implementar) as plataformas que suportarão o sistema de gestão de conteúdos, o trabalho colaborativo, e a comunicação entre os indivíduos do contexto. De referir que a motivação para o trabalho colaborativo é um aspeto que deverá começar a ser pensado e realizado desde a fase de planeamento, através do estabelecimento de uma comunicação regular entre os membros da equipa (e também com os colaboradores do contexto) através das plataformas, no sentido de se atingir os objetivos que já falamos nas anteriores secções deste capítulo.

As características individuais de cada contexto fazem com que as tarefas mencionadas acima não possam ser aplicadas de igual forma independentemente do contexto. Apesar do estabelecimento dos objetivos e benefícios da classificação ser uma tarefa aplicável a qualquer contexto, a calendarização das tarefas e a definição de recursos devem ser adequadas às características do contexto. Por exemplo, num contexto onde já se encontrem implementados sistemas de gestão de conteúdos e plataformas adequadas à comunicação e colaboração, não seria necessário implementar quaisquer outros *softwares* – apenas seria essencial selecionar quais as plataformas que suportariam o processo. O mesmo se passa com a definição da equipa e a calendarização de tarefas: quando o contexto já possui um departamento dedicado à gestão dos conteúdos organizacionais, não será tão evidente a importância dessas duas tarefas, uma vez que já estará subentendida nas responsabilidades profissionais desse departamento.

Infelizmente, o facto de as classificações não serem um aspeto muito relevante no planeamento dos projetos em que o INESC Porto está envolvido, tornou difícil a obtenção de informação relevante junto dos entrevistados no que diz respeito a este aspeto.

“[...] a direção do projeto define um núcleo de parceiros para o projeto, três ou quatro normalmente, que ficam encarregues das decisões acerca da organização de conteúdos do sistema, ou seja, qual informação que deve ser armazenada nesse sistema. Qual a informação pública, qual a informação privada, etc. Depois, alguém ficará encarregue dessa implementação e, por vez, nem sequer é alguém ligado diretamente ao projeto.”

Como a estrutura hierárquica de pastas utilizada ou obedece normalmente aos padrões definidas por uma entidade do projeto (no caso do *Plone*), ou está normalizada de acordo com as indicações do serviço de consultoria (no caso da *Dropbox*), não existe uma necessidade prioritária de decidir a forma como os conteúdos serão organizados em cada projeto. Outro fator que contribui para o não-desenvolvimento de classificação é o facto de as classificações utilizadas obedecerem a um modelo estrutural em hierarquia, e dos critérios que guiam a organização das pastas na hierarquia serem normalmente baseados na organização do próprio projeto – estes dois aspetos diminuem drasticamente as diferenças que poderiam existir entre as estruturas de pastas de diferentes projetos, uma vez que a organização dos projetos é, no geral, sempre similar. A juntar a este aspeto, também temos o aspeto relacionado com a própria conjectura em que os projetos são realizados, muito orientados para os deliverables e a sua avaliação. Se a necessidade de gestão do conhecimento estivesse mais presente durante o decorrer dos projetos, e consequentemente a necessidade de classificar os conteúdos de acordo com outras perspetivas mais específicas como assunto, área e serviço, veríamos, provavelmente, mais foco nas questões de gestão dos conteúdos e, particularmente, na sua classificação. O método que se apresenta de seguida pretende também ser um incentivo para a importância da gestão adequada dos conteúdos, mostrando que a colaboração num processo simples é suficiente para melhorar a gestão de conteúdos num contexto, e, por consequência, uma mais-valia para os processos e atividades deste.

5. Método para o desenvolvimento e gestão de classificações à medida de contextos específicos

A literatura revista e a informação recolhida e analisada ao longo deste trabalho forneceram uma base de conhecimento essencial para a construção de um método ágil e adequado ao desenvolvimento e gestão de classificações à medida de contextos específicos. As conclusões mais relevantes para a construção do método, retiradas da informação científica e daquela obtida junto dos profissionais, foram as seguintes:

- ✓ O âmbito do contexto empresarial deve regular o processo de desenvolvimento e gestão da classificação;
- ✓ A finalidade da implementação de uma classificação num contexto empresarial é a de ser útil para os processos organizacionais desse contexto específico, e não a de representar o conhecimento de forma universal;
- ✓ O processo de desenvolvimento e gestão de classificações deve ser levado a cabo de forma colaborativa e transparente, de forma a refletir as características do contexto onde foi elaborado – as classificações desenvolvidas em contextos empresariais específicos são relativas e enviesadas, e é este facto que as torna úteis para esse contexto;
- ✓ A classificação facetada parece ser a melhor estrutura dada a sua grande flexibilidade e multidimensionalidade, no entanto cada estrutura pode-se revelar útil em determinados contextos;
- ✓ O desenvolvimento colaborativo das classificações torna a classificação credível junto dos utilizadores e motiva a sua utilização, o que beneficia a partilha de conhecimento;
- ✓ As necessidades de recuperação de informação existentes no contexto são o aspeto que deve guiar o processo de desenvolvimento e gestão da classificação;
- ✓ A complexidade das necessidades de recuperação de informação do contexto são normalmente proporcionais à complexidade da futura classificação e do próprio processo de desenvolvimento e gestão desta;

- ✓ A adaptação da classificação à plataforma implementada tem uma influência tremenda na usabilidade e utilidade da classificação;
- ✓ A definição dos conceitos, termos, e possíveis facetas a representar na classificação é uma tarefa dependente das necessidades de recuperação de informação existentes no contexto, e como tal é difícil encontrar um processo normalizado aplicável a qualquer situação;
- ✓ Os mapas de conceitos são boas soluções para fomentar a colaboração neste processo, no entanto não são uma opção aplicável a todos os contextos;
- ✓ A utilização e gestão colaborativa e normalizada da classificação é tão ou mais importante que a colaboração durante o processo do seu desenvolvimento;
- ✓ A utilização de plataformas como meio de comunicação online é essencial para a gestão da classificação, uma vez que permite que os utilizadores deem o seu feedback, e discutam possíveis atualizações da classificação;
- ✓ O desenvolvimento de um guia de utilização da classificação é uma boa solução para motivar e garantir a utilização adequada e normalizada da classificação;
- ✓ Tal como em qualquer processo, a fase de planeamento é muito importante para se criarem bases para o sucesso do processo. Contudo, a importância das tarefas dessa fase variará devido à individualidade de cada contexto.

Apesar da lista acima não incluir toda a base de conhecimento útil na construção do método, é seguro dizer que nela estão representadas as principais conclusões obtidas na análise da literatura científica e nas entrevistas realizadas aos colaboradores do INESC Porto. Importante é também referir que a descrição do método a seguir apresentada conta já com algumas correções e adições que se mostraram como necessárias após a avaliação da primeira proposta do método (as alterações serão discutidas mais à frente, na secção de “avaliação da aplicação prática do método). Posto isto, de seguida é apresentado o principal resultado de todo o trabalho realizado nesta dissertação.

5.1 Descrição do método

O método para o desenvolvimento e gestão de classificações à medida de contextos específicos pretende auxiliar profissionais a desenvolverem e gerirem uma classificação para o seu contexto organizacional específico, através de uma série de tarefas agrupadas em etapas, e ordenadas sequencialmente de forma a representar a sua natureza processual. O método não exige uma aplicação rigorosa, em jeito de *checklist*, até porque nesse caso estar-se-ia a desconsiderar a importância da individualidade de cada contexto; pelo contrário, o método tenciona ser uma linha condutora para o desenvolvimento e gestão de classificações, ajudando os profissionais a adequarem da melhor forma o desenvolvimento e gestão da classificação às características do seu contexto específico.

5.1.1 Método: princípios

Apesar do método não exigir rigor na sua aplicação, esta deve obedecer e guiar-se segundo uma série de princípios gerais, cuja finalidade é o auxílio à criação de bases para a sua usabilidade e utilidade no contexto onde a classificação será implementada. Assim, os principais princípios a ter em conta no desenvolvimento e gestão de classificações são os seguintes:

- ✓ **Importância do contexto** – o âmbito do contexto empresarial deve regular o processo de desenvolvimento e gestão da classificação, para que esta possa refletir as suas características;
- ✓ **Foco no utilizador** - o processo de desenvolvimento da classificação deve ser sempre realizado tendo em mente as necessidades de recuperação de informação do futuro utilizador do sistema, no sentido de melhorar a usabilidade e utilidade da classificação;
- ✓ **Colaboração, comunicação e contribuição** - o desenvolvimento, gestão e utilização colaborativa das classificações torna a classificação credível e importante junto dos utilizadores e motiva a sua utilização;
- ✓ **Implementação tecnológica** - o software deve ser utilizado para potenciar o desenvolvimento, gestão e utilização da classificação. A classificação desenvolvida só será útil se for implementada adequadamente nos sistemas de gestão de conteúdos e/ou plataformas;
- ✓ **Utilidade** – o objetivo da classificação é ser útil para o contexto no qual vai ser implementado, e não de ser útil universalmente;
- ✓ **Relatividade** – o desenvolvimento e gestão de classificações não é um processo rígido de aplicação rigorosa, mas sim um processo dependente das

características do contexto, e que por isso exige criatividade e capacidade de brainstorming por parte dos envolvidos.

5.1.2 Método: processo

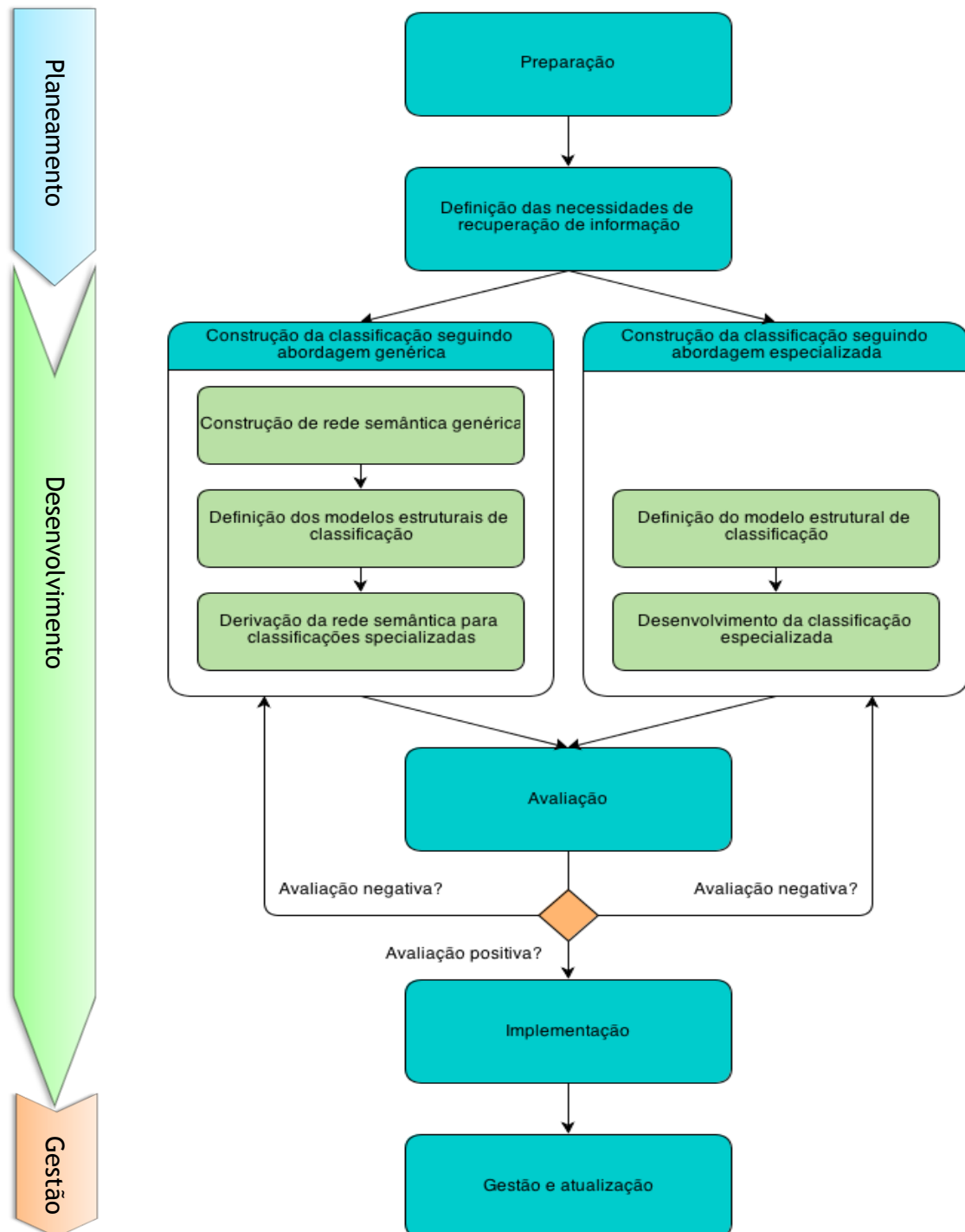


Ilustração 10 - Representação gráfica da proposta de método para o desenvolvimento e gestão de classificações à medida

Como é possível ver na representação gráfica do método (ilust. 10), o método encontra-se ordenado sequencialmente na vertical, de cima para baixo, e inclui várias etapas ou fases sequenciais. À esquerda, podemos ver que o método está organizado em três grandes fases:

1. **Planeamento** - nesta fase é efetuado o planeamento de todo o processo de desenvolvimento e gestão da classificação, assim como a análise das características do conteúdo e das necessidades de recuperação do contexto;
2. **Desenvolvimento** – nesta fase, a classificação sé desenvolvida, avaliada e implementado no contexto;
3. **Gestão** – esta fase inclui a gestão de atualização da classificação após a sua implementação e é, por natureza, uma fase contínua – é, essencialmente, a fase de gestão da qualidade da classificação.

Cada uma destas fases inclui uma ou várias etapas, que por sua vez incluem um conjunto de tarefas que serão descritas na próxima secção deste capítulo. Estas etapas estão representadas graficamente através dos retângulos azul-claros. Assim, na fase de *Planeamento*, foram incluídas as seguintes etapas:

- 1.1. **Preparação** – nesta etapa é feito o planeamento dos recursos necessários para as restantes etapas do processo, assim como a definição dos objetivos da classificação a desenvolver;
- 1.2. **Definição das necessidades de recuperação de informação** – nesta etapa é necessário analisar as características do conteúdo a classificar, assim como definir as necessidades de recuperação de informação do contexto, baseando-se tanto no futuro utilizador como na própria natureza dos conteúdos;

Seguidamente, a fase de *Desenvolvimento* inclui as seguintes etapas:

- 2.1. **Construção da classificação** – nesta etapa é construída a classificação, ou seja, é estruturado todo o sistema de classes de acordo com o modelo estrutural mais indicado para as necessidades de

recuperação de informação do contexto. Podem ser seguidas duas abordagens:

- a. Abordagem genérica* – esta abordagem implica o desenvolvimento colaborativo de uma rede semântica genérica, que posteriormente pode ser “desconstruída” para formar classificações especializadas para várias áreas específicas do contexto. É normalmente mais útil em contextos com dimensão reduzida, que não possuam um setor especializado em gestão de conteúdos, e que por isso queiram focar o trabalho colaborativo de uma equipa;
 - b. Abordagem especializada* – esta abordagem é mais direta que a genérica, uma vez que se constrói diretamente as classificações com os modelos estruturais indicados para satisfazer as necessidades recuperação definidas na fase de planeamento. É normalmente mais útil em contextos com grande dimensão, que possuam um setor especializado em gestão de conteúdos, que ficará encarregue do desenvolvimento e gestão da classificação ou classificações. Estes contextos normalmente não têm a disponibilidade de reunir uma equipa, com vários colaboradores de vários setores, para levarem a cabo um processo deste tipo. A abordagem especializada tem a desvantagem de requerer um novo processo de desenvolvimento de classificação caso surja uma nova necessidade de recuperação de informação;
- 2.2. Avaliação** – nesta etapa será avaliada a classificação, com a aplicação da mesma a uma amostra do conteúdo. Se avaliada positivamente, está pronta a ser implementada; se avaliada negativamente, deve ser corrigida e reavaliada até obter uma avaliação positiva;
- 2.3. Implementação** – nesta etapa a classificação é implementada nos sistemas de gestão de conteúdos, e é preparado a normalização do processo de utilização de classificação, a ser comunicado aos colaboradores do contexto.

Por último, na fase contínua de *Gestão*, temos a seguinte etapa:

3.1. Gestão e atualização – esta etapa contínua envolve a monitorização da utilização da classificação, através da análise do feedback dado pelos utilizadores da classificação, e a realização das atualizações e correções necessárias.

5.1.3 Método: indicações instrutórias

Tal como foi referido no capítulo anterior, nesta secção são explicadas as tarefas enquadradas em cada uma das etapas. Esta secção inclui, assim, as indicações instrutórias para a aplicação do método, e como tal, as tarefas são descritas com algum detalhe e contém, por vezes, algumas recomendações e/ou exemplos.

A lista de fases, etapas e tarefas está organizada de forma sequencial da seguinte forma (a ordem em que as tarefas estão organizadas na lista deverá ser a ordem de realização das mesmas):

1. FASE

1.1. Etapa

1.1.1. Conjunto de tarefas ou tarefa

1.1.1.1. Tarefa

Posto isto, é apresentada de seguida a lista completa das indicações instrutórias e recomendações:

1. PLANEAMENTO

1.1. Preparação

1.1.1. Definição dos recursos

1.1.1.1. *Definição da equipa encarregue do desenvolvimento e gestão da classificação*

1.1.1.2. *Seleção (e implementação) das plataformas colaborativas - será o meio de comunicação entre os membros da equipa, e mais tarde, também entre os colaboradores do contexto e a equipa (recomenda-se a utilização de fóruns e de wikis semânticas. A última particularmente permite não só a comunicação entre colaboradores mas também a gestão colaborativa da classificação diretamente nessa plataforma)*

1.1.2. Gestão do processo de desenvolvimento da classificação

1.1.2.1. *Calendarização das etapas e tarefas do processo*

1.1.2.2. *Motivação e incentivo ao trabalho colaborativo da equipa - incentivo à comunicação através das plataformas colaborativas, comunicação dos benefícios da classificação, registo e discussão dos problemas/dúvidas, entre outras tarefas que potenciam a coesão da equipa. Esta tarefa deverá, idealmente, ser contínua durante todo o processo.*

1.1.3. Estabelecimento dos objetivos e potenciais benefícios da classificação para a gestão e recuperação da informação no contexto

1.2. ***Definição das necessidades de recuperação de informação***

1.2.1. Análise dos conteúdos a serem classificados

1.2.1.1. *Definição do âmbito dos conteúdos – verificar se são conteúdos para disponibilizar ao público ou se apenas para gestão interna do projeto (deliverables e versões provisórias destes, informação técnica, entre outros)*

1.2.1.2. *Definição da utilidade futura dos conteúdos – verificar se são conteúdos com importância para o contexto em termos de gestão do conhecimento (a sua consulta futura será importante?)*

1.2.2. Definição das necessidades de recuperação de informação (de acordo com a natureza dos conteúdos e as necessidades de informação dos utilizadores do contexto) – um contexto terá, normalmente, mais que uma necessidade de recuperação de informação. Neste caso, poderão ser perfeitamente utilizados dois modelos estruturais na próxima fase, até porque todos eles se complementam. Por exemplo, o modelo por facetas deverá ser, normalmente, complementado por uma hierarquia de pastas, em que esta última representará a faceta mais importante dos conteúdos.

1.2.2.1. *Verificação da necessidade de recuperação de informação para gestão interna de acordo com a estrutura do contexto – recomenda-se o modelo estrutural em hierarquia (de pastas), de acordo com as atividades, processos e/ou tarefas realizadas no contexto. É importante ter em conta que se a quantidade de conteúdos a gerir no sistema se tornar muito elevada (superior a cerca de 1000 objetos informacionais), o tempo de recuperação*

poderá começar a tornar-se bastante longo, e começará a ser necessária a implementação de um modelo estrutural mais complexo, como a classificação facetada, que permite uma pesquisa eficaz por introdução de interrogações de pesquisa);

1.2.2.2. *Verificação da necessidade de recuperação de informação a disponibilizar ao público, por utilizadores externos ao contexto* – recomenda-se o modelo estrutural de classificação facetada, com a utilização das facetas necessárias para corresponder às possíveis necessidades do utilizador “típico” do tipo de conteúdos “públicos” que o contexto disponibiliza;

1.2.2.3. *Verificação da necessidade de recuperação de informação no âmbito da gestão do conhecimento para o futuro* – recomenda-se o modelo estrutural de classificação facetada, com a utilização das facetas necessárias para corresponder às possíveis necessidades do utilizador-colaborador do contexto (por exemplo, “área”, “serviço”, “tipo de documento”, “versão do documento”, entre muitas outras);

1.2.2.4. *Verificação da necessidade de recuperação de informação de acordo com necessidades pessoais e de manutenção de um ambiente incentivador da colaboração e participação na partilha de conhecimento* – recomenda-se o modelo estrutural da folksonomia (sistema de tagging), permitindo que o utilizador escolha permitindo que o utilizador escolha e atribua tags ao conteúdo de acordo com as suas perspetivas pessoais, contribuindo ativamente com a sua individualidade para a gestão colaborativa do conhecimento;

2. DESENVOLVIMENTO

2.1. Construção da classificação (optar por uma das duas abordagens)

ABORDAGEM GENÉRICA

2.1.1. Construção de rede semântica genérica - construção de um mapa de conceitos relacionados entre si, e relevantes para o contexto. Devem ser consideradas todas as perspetivas possíveis do contexto, de forma a se obter uma rede genérica. As típicas relações entre conceitos serão “x é um y” e “x tem y”. A colaboração e contribuição individual dos membros da

equipa são vitais nesta tarefa (recomenda-se a utilização das aplicação CmapTools ou ConceptME)

2.1.2. Definição dos modelos estruturais - escolha dos modelos estruturais de classificação mais adequados para satisfazer as necessidades de recuperação de informação definidas na etapa anterior (as recomendações estão indicadas na informação da etapa anterior)

2.1.3. Derivação da rede semântica para classificação(ões) especializada(s)

2.1.3.1. *Adaptação de parte da rede semântica genérica para a criação de uma classificação especializada utilizando o modelo estrutural adequado à necessidade de recuperação de informação existente no contexto* - ou seja, utilizar uma parte dos conceitos da rede semântica para criar uma classificação especializada, cujo modelo estrutural deverá ser selecionado de acordo com a necessidade de recuperação, encontrada na segunda etapa, que melhor se adapte a necessidade de recuperação de informação que origina essa classificação especializada. Normalmente, as relações da rede semântica definidas como “x é um y” resultam em classificações hierárquicas, e as relações “x tem y” resultam em classificação facetadas, onde “y” é uma faceta de “x”. Podem ser derivadas várias classificações a partir da rede semântica genérica.

2.1.3.2. *Atribuição de notações a cada classe da classificação especializada* (opcional)

ABORDAGEM ESPECIALIZADA

2.1.1. Definição do modelo estrutural - escolha do modelo estrutural de classificação mais adequados para satisfazer as necessidades de recuperação de informação definidas na etapa anterior (as recomendações estão indicadas na informação da etapa anterior). Pode ser necessário escolher dois modelos distintos, o que dará origem a duas classificações especializadas.

2.1.2. Construção da classificação especializada

2.1.2.1. *Definição dos conceitos relevantes para a classificação especializada, e escolha dos termos* (recomenda-se a consulta de tesouros e de outras classificações relacionadas com a área do contexto. Também pode ser útil a construção de um mapa de

conceitos pequeno para organizar o conhecimento e facilitar esta tarefa)

2.1.2.2. *Organização dos termos através da definição das relações* (na classificação facetada esta fase corresponde à criação de facetas (“tipo de documento”, “versão do documento”, “atividade”, “assunto”, etc.)

2.1.2.3. *Atribuição de notações a cada classe* (opcional)

2.2. Avaliação

2.2.1. Teste da classificação através da aplicação da mesma a uma amostra de conteúdo - classificar, em forma de teste, uma amostra do conteúdo típico a ser gerido no sistema

2.2.1.1. *Avaliação da funcionalidade (termos suficientes/insuficientes), usabilidade (facilidade ou dificuldade de aplicação da classificação) e utilidade (se é provável/improvável que a classificação venha a melhorar a recuperação dos conteúdos no contexto) da classificação*

2.2.1.2. *Identificação e registo dos problemas encontrados*

2.2.2. Possível correção/reformulação da classificação – alguns problemas comuns têm sintomas de relativamente fácil dedução: fraca usabilidade é sintoma de elevada complexidade ou profundidade na classificação; pouca utilidade é sintoma de erros na definição das necessidades de recuperação de informação ou na escolha do modelo estrutural mais adequado; fraca funcionalidade (termos insuficientes para classificar todo o tipo de conteúdos) é sintoma de pouca profundidade na classificação e de quantidade insuficiente de conceitos e relações relevantes

2.2.3. Reavaliação da classificação

2.3. Implementação

2.3.1. Seleção (e implementação) dos sistemas de gestão de conteúdos e/ou plataformas onde será utilizada a classificação

2.3.2. Adaptação da classificação às funcionalidades do sistema de conteúdos - o modelo estrutural de classificação pode ter que ser “modificado” quando aplicado às funcionalidades de classificação do sistema de conteúdos. Os sistemas de conteúdos não são, na maioria das vezes, capazes de suportar classificação mais complexas, como as facetadas. Por isso, recomenda-se

que se mantenha a classificação estabilizada externamente ao sistema (por exemplo, na plataforma colaborativa utilizada para gerir o processo de desenvolvimento da classificação), e que se faça a adaptação necessária aos sistemas onde a classificação será implementada. Normalmente, a adaptação passará pela criação de uma estrutura hierárquica de pastas (num modelo estrutural em hierarquia), ou pela atribuição de categorias a cada conteúdo de acordo com as várias facetas da classificação (no caso de uma classificação facetada, a faceta “principal” pode ser transposta para uma estrutura de pastas, e as restantes serem adaptadas ao sistema através da atribuição de categorias/etiquetas a cada conteúdo). No caso das folksonomias, as tags serão também atribuídas através das opções de atribuição de categorias. No entanto, caso os sistemas utilizados disponham de funcionalidades mais complexas que a simples atribuição de categorias ao conteúdo, poderão surgir novas formas de levar a cabo esta adaptação.

- 2.3.3. Modelação do processo de classificação do conteúdo e comunicação do mesmo aos colaboradores do contexto - criação de um guião com regras sobre como classificar o conteúdo no sistema, de forma a melhorar a normalização e usabilidade do processo
- 2.3.4. Classificação do conteúdo presente no sistema de gestão de conteúdos e/ou plataformas, de acordo a classificação

3. GESTÃO

3.1. Gestão e atualização

- 3.1.1. Análise do feedback dos colaboradores acerca da utilização da classificação - a utilização das plataformas colaborativas como meio de comunicação é muito importante nesta tarefa
 - 3.1.1.1. *Monitorização da utilidade da classificação – a classificação se satisfaz/não satisfaz as necessidades de recuperação de informação definidas na segunda etapa?*
 - 3.1.1.2. *Monitorização de possíveis novas necessidades de recuperação de informação*
 - 3.1.1.3. *Monitorização de incapacidades da classificação – a classificação não tem/ tem termos insuficientes?*

3.1.1.4. *Monitorização de erros na semântica da classificação - a classificação tem/não tem termos ou relações mal enquadrados e/ou definidos?*

3.1.2. Atualização da classificação

3.1.2.1. *Desenvolvimento de classificação com novo modelo estrutural para satisfazer necessidade de recuperação de informação emergente no contexto*

3.1.2.2. *Correção de erros na classificação atribuída pelos utilizadores aos conteúdos*

3.1.2.3. *Inclusão de novos termos e relações na classificação*

3.1.2.4. *Correção de erros na semântica da classificação*

5.2 Avaliação da aplicação prática do método

A avaliação da aplicação prática do método, descrito ao longo deste capítulo, foi realizada junto do serviço de consultoria do INESC Porto. Como já foi dito nesta dissertação, o serviço de consultoria do INESC Porto está encarregue de grande parte das iniciativas de melhoria da gestão dos conteúdos dos vários projetos em que a instituição está envolvida e, como tal, muitas das responsabilidades dos colaboradores deste serviço passam pela gestão da informação do contexto geral (INESC Porto), mas também pela ajuda na gestão dos conteúdos de cada contexto específico (cada projeto). Neste sentido, o serviço de consultoria assumiu-se como a melhor opção para avaliar uma proposta de método para desenvolvimento e gestão de classificações à medida, particularmente pela grande variedade de contextos distintos existentes dentro da instituição. Apesar de esta razão ter sido a principal no momento da escolha do contexto para avaliar a proposta de método, existiram outras, nomeadamente:

- ✓ A acessibilidade e disponibilidade do serviço de consultoria;
- ✓ O INESC Porto já havia sido um local de investigação durante o trabalho para a dissertação;
- ✓ Os colaboradores entrevistados durante a realização deste trabalho pertencem ao INESC Porto, e, dessa forma, os resultados deste trabalho pretendem retribuir a sua disponibilidade.

Então, a avaliação do método teve como foco os seguintes critérios:

- ✓ Funcionalidade – representa a capacidade do método em chegar a um resultado quando aplicado a uma situação, ou seja, se este é capaz de criar uma classificação para o contexto específico. Maior capacidade = maior funcionalidade;
- ✓ Eficiência – representa o período de tempo que a aplicação do método leva até à etapa de gestão, exclusive. Está diretamente à relação entre o tempo de aplicação e a complexidade da classificação. Melhor relação entre tempo de aplicação e complexidade da classificação = maior eficiência;
- ✓ Usabilidade – representa a facilidade de aplicação do método. É diretamente influenciada pela simplicidade ou complexidade deste. Maior facilidade = melhor usabilidade;
- ✓ Utilidade – representa a eficácia do resultado do método em satisfazer as necessidades de recuperação de informação definidas. Será mais difícil de avaliar devido ao período de tempo necessário para monitorizar a satisfação das necessidades no contexto. Além disso, pode ser condicionada por erros do utilizador do método durante a sua aplicação, particularmente na etapa de definição de necessidades.

A avaliação foi realizada com a “semi-participação” do autor da proposta do método. Isto porque, apesar da avaliação da aplicação do método ser da responsabilidade do serviço de consultoria, ela foi precedida por uma reunião informal, com presença do autor e dois colaboradores do serviço de consultoria, onde foi explicada e discutida a avaliação da aplicação prática do método no contexto do INESC Porto. A proposta do método utilizada na avaliação foi uma versão provisória, resumida e adaptada do método, que foi posteriormente melhorada com a ajuda da avaliação. Esta versão contava com a representação gráfica do método, com as indicações instrutórias, e também com algumas ajudas no que diz respeito ao tipo de resultados pretendidos em cada etapa (esta versão pode ser consultada no Anexo 2 desta dissertação). Nesta reunião foi já possível descortinar uma alteração muito importante para melhorar a compreensão e a aplicação prática do método, nomeadamente:

- ✓ A necessidade de instruções mais claras e detalhadas em cada uma das tarefas, uma vez que a aplicação do método pode tornar-se complicada para indivíduos sem competência na área da gestão da informação. Esta necessidade foi correspondida na proposta apresentada nesta dissertação através da inclusão de um guião simplificado complementar para aplicação do método (que pode ser consultado no Anexo 1), e através da inclusão de explicações mais detalhadas, exemplos e ainda recomendações em algumas das tarefas.

Posteriormente a esta reunião, a aplicação do método ficou totalmente na responsabilidade do serviço de consultoria, sem participação do autor. É importante referir, no entanto, que durante a reunião foi já “delineada” grande parte das opções a ser tomadas durante o processo, o que permitiu obter já grande parte da informação necessária para a avaliação do método. O registo da informação relevante para a avaliação foi efetuado pelos colaboradores num formulário com um esquema em tabela, e com uma organização similar à do próprio método, no sentido de tornar o registo de informação intuitivo. Desta forma, os responsáveis pela avaliação ficaram encarregues de registar os problemas encontrados e dificuldades experienciadas durante a realização de cada tarefa, quaisquer desvios feitos em relação às indicações do método e as razões/causas para os mesmos, os resultados obtidos em cada etapa e outras informações que parecessem relevantes para a avaliação do método. Como na reunião já tinham ficado definidas muitas das opções a ser tomadas durante o processo de aplicação do método, o registo da informação para cada etapa não foi totalmente completo.

Passando à avaliação de cada etapa da proposta de método, na fase de *Planeamento*:

- ✓ A etapa de *Preparação* foi realizada sem grandes problemas, no entanto foi necessário efetuar alguns desvios em relação ao método para a adaptação deste ao contexto ser bem-sucedida. A equipa responsável contou com 2-3 colaboradores do serviço de consultoria, a calendarização não foi levada a cabo porque a avaliação ficou circunscrita a um período de tempo curto (1 semana), assim como a motivação para a colaboração não foi absolutamente necessário, uma vez que a equipa partilha o mesmo

local físico de trabalho. O desvio mais importante, e que poderia ter repercussões negativas na futura gestão da classificação, foi o facto de não ter sido considerada nenhuma plataforma online para servir como meio de comunicação e colaboração – este desvio, no entanto, pode ser explicado pela período de tempo escasso que impede uma avaliação completa e profunda do método. Os objetivos definidos pelo serviço de consultoria para a classificação foram melhorar a gestão da informação para gestão interna de cada projeto, e melhorar a recuperação de informação para benefício da gestão do conhecimento e da sua utilidade em contextos futuros (durante e pós projeto). Estes dois objetivos são bastante interessantes na perspetiva em que permitem já descobrir as principais necessidades de recuperação de informação que serão fundamentais na próxima etapa. Outro aspeto importante nesta etapa está relacionado com a tarefa de análise do conteúdo: a natureza e finalidade desta tarefa tem uma relação mais próxima do cariz da segunda etapa do método, nomeadamente a etapa de *Definição das necessidades de recuperação de informação*. Como essa tarefa estava, na versão provisória do método utilizada na avaliação, incluída na etapa de *Preparação*, optou-se por transpor a mesma para a segunda etapa desta fase após uma posterior análise do método;

- ✓ A etapa de *Definição das necessidades de recuperação de informação* foi realizada com base no que havia sido definido como objetivos da classificação na etapa anterior, o que permitiu que as tarefas da segunda etapa fossem concluídas com relativa facilidade. No que diz respeito à análise do conteúdo, os conteúdos foram definidos como sendo “técnicos” e “de gestão do projeto”, o que indica, logicamente, que os conteúdos serão relevantes para o contexto interno de cada projeto e, segundo o que foi definido nos objetivos da classificação, também para a questão da gestão do conhecimento dentro do contexto geral do INESC Porto. Os objetivos e a análise dos conteúdos permitiram uma transição simples para a definição de necessidades de recuperação de informação. Então, e de acordo com o serviço de consultoria, existem duas principais necessidades no contexto geral do INESC Porto: recuperação de informação para gestão interna de cada projeto e recuperação de

informação no âmbito da gestão do conhecimento para o futuro, num contexto mais geral onde estão incluídos os vários projetos de consultoria em que o INESC Porto está e esteve envolvido.

Passando à fase de *Desenvolvimento*:

- ✓ Os colaboradores do serviço de consultoria optaram pela abordagem especializada na etapa de *Desenvolvimento da classificação* por duas razões. Primeiro, a área de Consultoria já dispõe de uma classificação hierárquica, na forma de uma estrutura de pastas normalizada, utilizada nos projetos de consultoria, organizada de acordo com o contexto da gestão de projetos “e orientada aos resultados e às tarefas a executar para atingir esses resultados”; logo, só necessitam de um complemento para satisfazer a necessidade de recuperação no âmbito da gestão do conhecimento. Segundo, as restrições em termos de tempo diminuía a viabilidade da abordagem genérica, que necessita de um maior período de tempo para ser realizada convenientemente. Passando à definição dos modelos estruturais, segundo o que foi possível perceber, as recomendações incluídas no método tornarão essa tarefa simples, sendo apenas necessário fazer a correspondência entre as necessidades de recuperação de informação e os modelos estruturais de classificação recomendados para cada caso. Como o contexto já dispunha de uma classificação hierárquica para corresponder à necessidade de recuperação para gestão interna do projeto, foi intuitivo que o foco do processo passaria para o desenvolvimento de uma classificação facetada para corresponder à necessidade de recuperação de informação no âmbito da gestão do conhecimento para o futuro. A construção da classificação especializada acabou, assim, por se transformar num esforço de definição de facetas adequadas às características partilhadas por todos os projetos do contexto, de forma a construir uma classificação facetada capaz de satisfazer utilizadores de vários tipos de contextos mais específicos (cada projeto). Devido às restrições de tempo (não só o período de tempo para avaliação da proposta de método era curto, mas também o serviço de consultoria conta com várias responsabilidades dentro do contexto), a

equipa optou por basear a classificação facetada numa outra classificação anteriormente desenvolvida, pela aluna Diana Vieira na sua dissertação de Mestrado em Ciência da Informação¹², para o contexto do INESC Porto. Com esta base, a equipa definiu as seguintes facetas:

- Fonte do conteúdo;
- Domínio de intervenção;
- Tipo de indústria;
- Área científica;
- Tipologia documental;
- Formato do documento;
- Língua.

O trabalho de Diana Vieira inclui, ainda, vários termos para cada faceta, que se mostram importantes para normalizar e limitar os termos que podem ser atribuídos a cada conteúdo. Isto permitiu, assim, validar a solidez das facetas dentro do contexto. No que diz respeito à notação das classes, a equipa optou por manter apenas a notação que já mantinha para documentos individuais, nomeadamente “[SiglaProjeto].tp.ff.pp.parceiro.aaaammdd.ddddd.VD.d”, em que:

- Tp: tipo de documento;
 - Ff: fase do projeto;
 - Pp: processo;
 - Parceiro;
 - Aaaammdd: data;
 - VD.d: versão.
- ✓ A etapa de *Avaliação* também constitui outro desvio em relação ao método, uma vez que a equipa baseou a sua classificação no trabalho já validado de Diana Vieira;
 - ✓ As opções a tomar durante a etapa de *Implementação* foram discutidas, na sua maioria, durante a reunião, devido às restrições de tempo que não permitiram levar a cabo todas as tarefas do processo num âmbito prático. Apesar de tudo, a eficácia do método depende sobretudo das decisões

¹² Vieira, Diana (2012). “Definição de uma Estratégia para a gestão de conteúdos: o caso de estudo da área de consultoria da unidade de Engenharia de Sistemas de Produção do INESC TEC”. Dissertação de mestrado, Faculdade de Engenharia da Universidade do Porto.

tomadas em cada etapa e tarefa, e não tanto da aplicação prática dessas decisões. Passando então às tarefas desta etapa, a equipa escolheu a *Dropbox* como a plataforma onde seria implementada a classificação, o que era expectável dado que era nesta onde estavam armazenados os vários conteúdos de cada projeto. Em relação à adaptação da classificação à *Dropbox*, decidiu-se que seria efetuada através da atribuição de etiquetas a cada conteúdo, tal como recomendado na proposta de método. Esta adaptação exige uma modelação do processo de classificação adequada para o utilizador compreender de que forma deve classificar o conteúdo – para isso, a equipa decidiu que a melhor forma seria normalizar o processo de forma semelhante à sua notação standard para documentos, ou seja, indicar as facetas e os termos possíveis num guião, assim como o processo que o utilizador teria de seguir para introduzir as várias etiquetas corretas em cada conteúdo. Como foi referido, as restrições de tempo impediram que estas decisões fossem aplicadas na prática, e como tal também não foi possível concluir a tarefa de classificação de conteúdo armazenado no sistema (que levaria um grande período de tempo a completar dada a enorme quantidade de conteúdos armazenados na plataforma *Dropbox*).

No que diz respeito à avaliação da fase de *Gestão*, é impossível realizar um verdadeiro teste, uma vez que essa fase é baseada na monitorização da utilização da classificação e na sua possível atualização, e esse tipo de tarefas requer um período de tempo alargado de utilização da classificação em contexto real de trabalho. Como tal, optou-se por deixar em segundo plano essa fase.

A avaliação realizada no contexto do serviço de consultoria do INESC Porto sugeriu que o método é exaustivo no que diz respeito às etapas e tarefas necessárias para a criação de uma classificação adequada ao contexto. Neste sentido, é possível afirmar com alguma certeza que o método possui as funcionalidades necessárias para atingir o resultado que pretende, uma vez que no teste realizado no contexto complexo da área de consultoria do INESC Porto foi possível alcançar o resultado pretendido. A questão da **funcionalidade** do método não é tão dependente da quantidade de testes realizados, tendo em conta que apenas está relacionada com a capacidade do método

ajudar a produzir um resultado, independentemente da qualidade deste. Neste sentido, como o método foi suficiente para produzir um resultado no contexto desta avaliação, é relativamente seguro afirmar que o mesmo se passaria noutros contextos. Em termos de **eficiência**, é arriscado retirar conclusões gerais, considerando que não foi possível completar na prática todas as tarefas da etapa de implementação, e que a avaliação deste aspeto requer uma quantidade considerável de testes. A equipa encarregue da avaliação conseguiu tomar as decisões necessárias num breve período de tempo, o que indica que a eficiência do método dependerá sobretudo do tempo necessário para a implementação prática dessas decisões (o que não foi possível averiguar, em parte). Mas, mais uma vez, é importante ter em conta que a eficiência do método é muito dependente das características do contexto – por exemplo, no contexto onde foi realizada a avaliação, a quantidade de conteúdos armazenados na plataforma sugeria um período de implementação bastante alargado, tendo em conta que o método prevê a classificação do conteúdo já existente; por oposição, se a avaliação fosse realizada pré-implementação da própria plataforma, esta tarefa nem seria necessária, uma vez que não existiam ainda conteúdos armazenados. Um aspeto positivo que foi possível retirar da avaliação é a inexistência de etapas totalmente desnecessárias ou redundantes no método – a equipa necessitou de todas as etapas do método para alcançar os resultados pretendidos, e não teve dúvidas sobre a pertinência de nenhuma delas.

No que diz respeito à **usabilidade** do método, a exaustividade, que torna o método funcional, acarreta, no entanto, um aspeto negativo, nomeadamente a possibilidade do método conter alguma complexidade quando utilizado por indivíduos sem competências adequadas na área da gestão da informação. Este problema é em parte corrigido pela adição de um guião simplificado (Anexo 1), tal como já foi mencionado. Mais uma vez, o teste do método noutros contextos distintos, com outro tipo de colaboradores com diferentes competências, permitiria uma avaliação mais adequada deste aspeto, que não foi possível devido às restrições de tempo e de acessibilidade. Apesar de tudo, a adição do guia simplificado torna o método mais acessível para indivíduos com menos competência na área da gestão da informação, sendo que os indivíduos com mais conhecimento nessa área podem consultar e utilizar a versão normal, mais complexa e detalhada. Por último, a avaliação da **utilidade** do método também sofreu com as restrições de tempo e de acessibilidade a outros contextos, que não permitirem que fosse possível implementar a classificação na

prática e monitorizar a sua utilização no contexto da área de consultoria do INESC Porto, e também que os testes de avaliação do método pudessem ser realizados noutros contextos com características distintas. Assim, e infelizmente, não foi possível retirar conclusões relevantes quanto a este aspeto.

Concluindo a avaliação, verificou-se que a proposta de método tem potencial para ser bastante útil na tarefa a que se propõe, no entanto a obtenção de certezas neste campo requer um maior período de avaliação, e uma maior quantidade de testes noutro tipo de contextos que se disponham a implementar os resultados deste método no seu contexto real de trabalho. Foi impossível obter este tipo de condições de avaliação no âmbito desta dissertação de mestrado, e, infelizmente, o método acaba por ficar prejudicado neste aspeto.

6. Conclusões e perspectivas de desenvolvimento

A crescente valorização da informação como ativo organizacional tem feito da Gestão de Conteúdos Empresariais (ECM) uma área cada vez mais importante para a indústria das tecnologias de informação e comunicação e para o mundo empresarial. Como vimos ao longo deste trabalho, as organizações têm procurado soluções tecnológicas que melhor se adaptem aos seus contextos organizacionais específicos, no entanto essa adaptação passa também pelo desenvolvimento de melhores métodos de organização da informação. As classificações mostram-se um importante instrumento nesta tarefa, ao ser uma mais-valia para a organização e recuperação dos conteúdos organizacionais. Apesar disso, o desenvolvimento de classificações adequadas às características específicas de cada contexto pode tornar-se uma atividade complicada para os profissionais – desde aspeto surge a necessidade de definição de um método que ajuda no desenvolvimento e na gestão de classificação à medida de contextos específicos, necessidade que este trabalho pretende ajudar a suprimir.

O método desenvolvido neste trabalho procurou não só ter uma natureza fundamentalmente prática, mas também basear os seus princípios e o seu processo no conhecimento científico proveniente da área da Ciência da Informação (CI) e nos princípios da área profissional da ECM. A revisão da literatura científica permitiu perceber que existe uma divisão de visões dentro da área da CI, basicamente fundamentadas em duas perspectivas da classificação de informação com princípios baseados em naturezas distintas: primeiro, a visão mais tradicional, onde a classificação da informação assenta nos princípios técnicos clássicos, e que privilegia o rigor da definição e dos processos e a representação universal da informação; em segundo, a visão influenciada pela prática profissional, que defende a utilidade organizacional da classificação, e que privilegia a utilização das classificações na resolução de problemas específicos de gestão da informação e como instrumento de apoio às atividades e processos específicos de cada contexto, para além da influência da relatividade no desenvolvimento de classificações, causada pela individualidade dos contextos. A perspectiva defendida neste trabalho, e que serviu de base científica para o método desenvolvido, centra-se sobretudo na segunda visão, centrada na resolução de problemas de organização de informação em contextos organizacionais específicos, mas

que no entanto não descarta a importância dos princípios tradicionais da classificação que são extremamente úteis na manutenção da coerência e coesão estrutural das classificações. Neste sentido, é possível concluir que o primeiro objetivo deste trabalho - *definir uma visão que guie a criação de classificações à medida de forma eficaz* – foi alcançado, na medida em que foi possível definir uma perspetiva única que se adaptou de forma bem-sucedida à pretendida aplicabilidade prática do método a contextos organizacionais específicos. Esta perspetiva ajustou-se ao tipo de estratégias que as organizações têm procurado no âmbito da ECM, onde a influência direta do contexto nas tomadas de decisões a este nível é visível, e a integração dos sistemas de gestão de conteúdos é a meta a alcançar pela maioria das organizações.

Numa investigação enquadrada na metodologia de *design science* e com um cariz exploratório, procurou-se complementar o conhecimento adquirido na revisão de literatura científica e do estado-de-arte com a experiência prática dos profissionais diariamente envolvidos em atividades onde a recuperação de informação é fulcral. As entrevistas realizadas com os colaboradores do INESC Porto permitiram encontrar padrões de utilização e organização de informação, assim como definir os vários tipos de necessidades de recuperação de informação que podem existir num contexto organizacional. Em conjugação com o conhecimento proveniente da revisão de literatura, foi possível deduzir que estas necessidades poderiam ser satisfeitas com a utilização de classificações com modelos estruturais adequados. Estes padrões de classificação de conteúdos adequados a tipos de necessidades de recuperação de informação específicas foi a base principal para a edificação do método, complementado por um processo que privilegia as atividades colaborativas como garantia de um resultado mais adequado e útil para o contexto específico. Neste sentido, é possível concluir que o segundo e terceiro objetivos - *identificar padrões para a organização e classificação de conteúdos e especificar processos colaborativos para o desenvolvimento e gestão da meta-informação e da classificação*, respetivamente -, foram suficientemente alcançados, na medida em que contribuíram para a construção de um método baseado num processo colaborativo onde os padrões de classificação de informação são o aspeto central.

A avaliação do método sofreu com as restrições de tempo e de acessibilidade a diferentes contextos, e como tal apenas foi possível comprovar que a proposta de

método possui potencial para constituir um importante auxílio para os profissionais no desenvolvimento e gestão de classificações.

O trabalho desenvolvido e relatado ao longo desta dissertação permitiu compreender a relevância que a organização e classificação adequada dos conteúdos poderá ter na recuperação da informação e, conseqüentemente, na própria eficiência dos processos organizacionais. Particularmente no que diz respeito à gestão do conhecimento organizacional, é fundamental que as organizações considerem o desenvolvimento de uma classificação adequada ao seu contexto individual como um investimento na melhoria da sua eficácia e eficiência. Um dos grandes problemas parece estar na própria tarefa de atribuição de classificação a conteúdos, que parece ser muitas vezes considerado como um aspeto acessório no momento de armazenamento do conteúdo, uma vez que o seu benefício não é imediato (só será “visível” no momento futuro da pesquisa), mas a sua contrapartida é – o facto de ser necessário gastar recursos de tempo a atribuir adequadamente a classificação. Quanto a este aspeto, cabe às organizações procurar formas de motivar e normalizar a utilização das suas classificações, sendo que o método desenvolvido ao longo deste trabalho procura focar essa necessidade (particularmente na etapa de *implementação*). Outro aspeto importante está relacionado com o tipo de modelos estruturais utilizados nas classificações: a utilização das estruturas hierárquicas de pastas baseadas na própria estruturação do contexto, apesar de ser suficientemente eficaz no tempo presente desse contexto específico (por exemplo um projeto), deixa de o ser quando se “entra” no campo da gestão do conhecimento para o futuro. A gestão do conhecimento requer um tipo de pesquisa cujas necessidades não são satisfeitas pelo modelo hierárquico – a hierarquia de pastas garante apenas uma pesquisa por navegação de acordo com uma única perspetiva, o que se mostra limitado no âmbito da gestão do conhecimento onde é importante pesquisar sob várias perspetivas numa vasta quantidade de conteúdos, que tornam a pesquisa por navegação demasiado ineficiente ao ponto de não ser eficaz nem amigável para o utilizador.

Apesar da classificação facetada se apresentar como uma solução eficaz para o problema da classificação no âmbito da gestão do conhecimento, ela traz consigo, infelizmente, outro problema. O modelo facetado, estruturalmente mais complexo, torna o desenvolvimento da classificação previsivelmente mais difícil para profissionais que não possuam as competências adequadas no âmbito da gestão da informação.

Apesar da proposta de método apresentada tentar colmatar este problema com o guia simplificado e as indicações instrutórias detalhadas, continua a ficar do lado dos profissionais as grandes decisões no momento da definição de facetas e de termos adequados para as mesmas – este aspeto é fulcral na futura utilidade da classificação e, infelizmente, continua a ser bastante difícil de “normalizar” e “guiar” num método como aquele proposto nesta dissertação. Outro problema reside na adaptação deste tipo de modelos estruturais a plataformas e *software* que não dispõe de opções mais avançadas no que diz respeito à classificação da informação. Como muitas organizações privilegiam, normalmente, a praticabilidade à complexidade, acabamos por verificar muitas plataformas a terem funções de sistemas de gestão de conteúdos quando na sua essência não foram criadas para isso (como por exemplo, o caso da *Dropbox*). Apesar da opção das organizações ser legítima e sensata - considerando a importância da acessibilidade e usabilidade dos *softwares* para os colaboradores -, ela acaba por tornar a adaptação das classificações às plataformas uma tarefa não muito imediata e que poderá se tornar demasiado complexa para profissionais sem as competências adequadas a este nível (e mesmo para quem as tem, este tipo de adaptação também requer algum know-how que dependerá da experiência e mesmo da capacidade de “improviso” do profissional). Estes dois problemas mostram mais uma vez a importância do papel do profissional da área da CI nas organizações, cuja abrangência de competências se adequam não só à resolução de problemas relacionados diretamente com a gestão do conhecimento, mas também com o aspeto da adaptação de instrumentos de organização de informação, como são as classificações.

O potencial do método desenvolvido ao longo desta dissertação é suficiente para se considerar que os objetivos principais deste trabalho foram alcançados, apesar da avaliação da proposta não ter sido tão extensiva e exaustiva quanto desejável. Ao longo do seu desenvolvimento foi possível perceber que a natureza individual de cada contexto na construção de classificação detém uma importância muito maior do que a que seria adequada para metodizar o processo de uma forma mais detalhada e normalizada. Mesmo assim, foi possível, através do conhecimento obtido na revisão de literatura e nas entrevistas aos profissionais, compor um método com uma funcionalidade e usabilidade adequada para ser utilizado como suporte para um processo de desenvolvimento e gestão de classificações. Mais uma vez, é importante notar que uma avaliação totalmente eficaz do método só seria possível com uma

quantidade de testes elevada em contextos reais de trabalho, e durante um período de tempo considerável, de forma a se poder avaliar não só o desenvolvimento das classificações, mas também monitorizar a sua utilização e avaliar a sua verdadeira utilidade para o contexto. A variedade de competências, características dos contextos, necessidades de recuperação de informação, e plataformas tecnológicas, faz com que seja necessário uma amostra maior de contextos para se poder tirar conclusões mais seguras e generalizadas no que diz respeito a métodos e outros resultados deste tipo. No entanto, uma avaliação com esta envergadura não foi possível durante o trabalho desta dissertação e, com base nisso, só é seguro e possível assumir o potencial desta proposta de método como suporte aos profissionais no processo de desenvolvimento e gestão de classificações à medida de contextos específicos.

Com base nestas considerações finais, é possível definir algumas linhas orientadores de possível trabalho futuro no âmbito dos objetivos desta dissertação, nomeadamente:

1. Procurar uma maior normalização da proposta de método no que diz respeito às questões relacionadas com a definição de facetas e termos a incluir na classificação;
2. Desenvolver o método no que diz respeito à colaboração no processo de desenvolvimento e gestão de classificação, com a inclusão de mais recomendações a esse nível;
3. Incluir propostas de tipos de plataformas de gestão de conteúdos adequadas a cada tipo de necessidade de recuperação, e opções diversas a nível de plataformas para o suporte ao trabalho colaborativo;
4. Realizar uma avaliação mais aprofundada do método, com a realização de vários testes em distintos contextos organizacionais reais.

A primeira perspetiva de trabalho futuro tem o intuito de melhorar a usabilidade do método junto dos profissionais com menos competência na área da gestão da informação, algo que foi possível constatar durante a avaliação da proposta de método. A maior normalização desta parte da proposta faria com que estes profissionais pudessem ter mais apoio durante esta fase, que é por natureza bastante dependente das características individuais do contexto.

A segunda perspetiva de trabalho futuro surge por razões muito semelhantes à primeira, na medida em que a motivação para a colaboração também é um aspeto que recai muito nas próprias características individuais dos colaboradores do contexto, e das próprias capacidades de liderança dentro do contexto. Mais uma vez, a melhoria deste aspeto seria ajudaria a retirar algum do peso desta atividade dos ombros da equipa encarregue do processo de desenvolvimento da classificação.

A terceira perspetiva pretende preencher uma lacuna deste trabalho, nomeadamente um dos sub-objetivos que previa a identificação e caracterização de tipos de plataformas de gestão de conteúdos adequadas a cada tipo de contexto ou necessidade de recuperação de informação. Este objetivo não foi alcançado devido à falta de tempo para análise adequada e utilização adequadas das funcionalidades das várias plataformas existentes no mercado de forma a as poder corresponder a cada tipo de contexto. A acessibilidade às plataformas também constituiu outro obstáculo a este objetivo. A melhoria deste aspeto traria uma nova dimensão, mais tecnológica, à proposta de método, o que seria uma excelente mais-valia.

Por último, a avaliação mais aprofundada da proposta de método traria, como já foi dito, uma maior validação do seu potencial. A realização de testes prolongados em contextos organizacionais variados e a monitorização da utilização das classificações resultantes da aplicação da proposta permitiriam não só avaliar mais adequadamente o método, mas também, efetuar reformulações que poderiam trazer melhorias significativas à aplicabilidade prática do método em contexto real de trabalho. No entanto, o potencial demonstrado pelo método desenvolvido é suficiente para afirmar que o trabalho desenvolvido ao longo da dissertação satisfaz os objetivos definidos inicialmente. Desse modo, pode-se concluir que esta dissertação apresenta contribuições para o conhecimento científico da ECM, ao providenciar uma proposta de método para o desenvolvimento e gestão de classificações à medida de contextos específicos, baseado numa perspetiva assente numa perspetiva científica, mas complementemente orientada para a prática profissional e para a utilidade da classificação para os processos organizacionais.

Referências bibliográficas

1. AIIM (2011). “State of the ECM Industry: how well is it meeting business needs”. *AIIM*, acessado a 5 de dezembro de 2012, <http://www.aiim.org/Research-and-Publications/Research/Industry-Watch/State-of-the-ECM-Industry-2011>
2. AIIM (2012). “What is Enterprise Content Management (ECM)?”. *AIIM*, acessado a 20 de novembro de 2012, <http://www.aiim.org/What-is-ECM-Enterprise-Content-Management>.
3. AIIM (2012). “What is Enterprise Search?”. *AIIM*, acessado a 20 de dezembro de 2012, <http://www.aiim.org/What-is-Enterprise-Search>.
4. AIIM Training (2012). “How to Develop Taxonomies to Support Navigation , Information Discovery , and Findability”. *AIIM*, acessado a 20 de novembro de 2012, <http://pages2.aiim.org/rs/aiim/images/Taxonomy-Briefing-2012-v01.pdf>
5. Alalwan, J. A., Wesroffer, H. R. (2012), “Enterprise content management research: a comprehensive review”. *Journal of Enterprise Information Management*, **25**(5), 441 – 461, acessado a 25 de outubro de 2012, <http://www.emeraldinsight.com/journals.htm?articleid=17054879>.
6. Alfresco (2012). “Tour: learn about Alfresco”. Alfresco Software Inc., acessado a 20 de dezembro de 2012, <http://www.alfresco.com/tour>.
7. Amorim, Fabiana e Tomaél, Maria (2011) “Gestão da informação e gestão do conhecimento na prática organizacional : análise de estudos de casos.” *Revista Digital De Biblioteconomia e Ciência Da Informação*, 8(2), 1–22, acessado a 20 de novembro de 2012, <http://www.sbu.unicamp.br/seer/ojs/index.php/rbci/article/view/465/316>.
8. Anderson, J.; Pérez-Carballo, J. (2001a). “The nature of indexing: how humans and machines analyze messages and texts for retrieval. Part I: Research, and the nature of human indexing.” *Information Processing and Management*, **37**(2), 231-254, acessado a 9 de janeiro de 2013, <http://comminfo.rutgers.edu/~muresan/IR/Docs/Articles/artAnderson2001a.pdf>.
9. Anderson, J.; Pérez-Carballo, J. (2001b). The nature of indexing: how humans and machines analyze messages and texts for retrieval. Part II: Machine indexing, and the allocation of human versus machine effort. *Information Processing and*

- Management*, **37**(2), 231-254, acessado a 9 de janeiro de 2013,
<http://www.calstatela.edu/faculty/jperezc/Publications/artAnderson2001b.pdf>.
10. Arshad, N. I., Bosua, R., Milton, S. K. (2010) "Facilitating Information Sharing in Organizations using Electronic Content Management Systems (ECMS): Towards a Model". *21st Australasian Conference on Information Systems*, Austrália, 2010, acessado a 25 de outubro de 2012, <http://aisel.aisnet.org/acis2010/45/>.
 11. Arshad, Noreen; Bosua, Rachelle; Milton, Simon K. (2010). "Facilitating Information Sharing in Organizations Using Electronic Content Management Systems (ECMS): Towards a Model Management Systems (ECMS): Towards a Model." Em *21st Australasian Conference on Information Systems*, acessado a 20 de novembro de 2012, <http://aisel.aisnet.org/acis2010/45/>.
 12. Bilski, Adrian (2011). "A Review of Artificial Intelligence Algorithms in Document Classification." *Journal of Electronics and Telecommunications*, **57** (3), 263–270.
<http://www.degruyter.com/view/j/eletel.2011.57.issue-3/v10177-011-0035-6/v10177-011-0035-6.xml>.
 13. Bonifacio, M.; Bouquet, P.; aRoberta Cuel (2002). "The role of classification(s) in distributed knowledge management."
 14. Brodtkin, J. (2008). "Autonomy, Endeca rate among top enterprise search vendors". *Network World*, acessado a 20 de dezembro de 2012,
<http://www.networkworld.com/news/2008/070208-top-search.html?page=2>.
 15. Broughton, Vanda (2006). "The Need for a Faceted Classification as the Basis of All Methods of Information Retrieval". *Aslib Proceedings* 58 (1/2): 49–72, acessado a 13 de fevereiro de 2013,
<http://www.emeraldinsight.com/10.1108/00012530610648671>.
 16. Broughton, Vanda e Slavic, Aida (2007). "Building a Faceted Classification for the Humanities: Principles and Procedures." *Journal of Documentation* 63 (5): 727–754, acessado a 20 de dezembro de 2012,
<http://www.emeraldinsight.com/10.1108/00220410710827772> .
 17. Broughton, Vanda, et al. (2011). "Brian Vickery and the Classification Research Group : the Legacy of Faceted Classification." Em *ISKO UK Conference 2011*.
 18. Buchanan, Bruce G. (2005). "A (Very) Brief History of Artificial Intelligence". *AI Magazine* **26**(4), 53–60, acessado a 25 de março de 2013,
<http://www.aaai.org/ojs/index.php/aimagazine/article/view/1848/1746>.

19. Chandrasekaran, B, Josephson, John R, and Benjamins , V. Richard.(1999). “What Are Ontologies , and Why Do We Need Them ?”. *IEEE Intelligent Systems*, 20–26, acessado a 25 de março de 2013,
<http://www.csee.umbc.edu/courses/771/papers/chandrasekaranetal99.pdf>.
20. Chandrasekaran, B. (1988). “From Numbers to Symbols to Knowledge Structures: Artificial Intelligence Perspectives on the Classification Task.” *IEEM Transactions on Systems, Man, and Cybernetics*, **18** (3), 415-424,
<http://home.cc.gatech.edu/dil/uploads/num-sym-knowledge.pdf>.
21. Chen, Jianqing et al. (2011). “Moderated Online Communities and Quality of User-Generated Content.” *Journal of Management Information Systems* **28**(2), 237–268. doi: 10.2753/MIS0742-1222280209.
22. Clancey, William J. (1984). Classification Problem Solving. *University of Stanford*.
23. CMSCritic (2012). “List of Enterprise Content Management Systems”. *CMSCritic*, acessado a 20 de dezembro de 2012, <http://www.cmscritic.com/resource-lists/ecm-list/>.
24. Costa, Luís et al. (2012). “ConceptME: Gestão colaborativa de modelos conceptuais”. Acessado a 30 de março de 2013,
http://www3.dsi.uminho.pt/CAPSI2012/CD/submissions/capsi2012_submission_34.pdf.
25. Davis, R., Shrobe, H., Szolovits, P. (1993). “What Is a Knowledge Representation?”. *AI Magazine* **14**(1), acessado a 20 de novembro de 2012,
<http://www.aaai.org/ojs/index.php/aimagazine/article/view/1029/947>.
26. Davis, Randall; Shrobe, Howard; e Szolovits, Peter (1993). “What Is a Knowledge Representation ?” **14** (1), 17–33.
27. Delphi Group (2004). “Information Intelligence: Content Classification and the Enterprise Taxonomy Practice.” Delphi Group, acessado a 25 de março de 2013,
<http://www.delphigroup.com/whitepapers/pdf/20040601-taxonomy-WP.pdf>.
28. DeltCI (2008). “Gestão da informação”. *DeltCI*, acessado a 20 de dezembro de 2012,
<http://www.ccje.ufes.br/arquivologia/deltci/def.asp?cod=41>.
29. Detlor, B. (2010). “Information management”. *International Journal of Information Management*, 30, 103-108, acessado a 25 de outubro de 2012,
<http://www.sciencedirect.com/science/article/pii/S0268401209001510#>.
30. Doculabs (2007). “Ten Questions on Taxonomy.”

31. Duffy, J. (2001), "The tools and technologies needed for knowledge management". *Information Management Journal* **35**(1), acessido a 7 de janeiro de 2013, <http://www.thefreelibrary.com/The+Tools+and+Technologies+Needed+for+Knowledge+Management.-a079742902>.
32. Evelson, B. (2011). "Top 10 business intelligence predictions for 2012". *Forrester*, acessido a 20 de dezembro de 2012, http://blogs.forrester.com/boris_evelson/11-11-15-top_10_business_intelligence_predictions_for_2012.
33. Gartner (2012). "Enterprise Content Management (ECM)". Gartner, acessido a 20 de dezembro de 2012, <http://www.gartner.com/it-glossary/enterprise-content-management-ecm/>.
34. Gnoli, Claudio (2011). "Vickery 's Late Ideas on Classification by Phenomena and Activities." In *ISKO UK Conference 2011*, acessido a 23 março de 2013, <http://www.iskouk.org/conf2011/papers/gnoli.pdf>.
35. Gruber, Thomas (1993). "Toward Principles for the Design of Ontologies Used for Knowledge Sharing". *International Journal Human-Computer Studies* **43**, 907-928, acessido a 25 de março de 2013, <http://tomgruber.org/writing/onto-design.pdf>.
36. Guarino, Nicola (1995). "Formal Ontology Conceptual Analysis and Knowledge Representation." *Human-Computer Studies* **43**.
37. Hevner, Alan R et al. (2004). "Design Science in Information Systems Research". *MIS Quarterly* **28**(1), 75-105, acessido a 12 de janeiro de 2013, <http://www.csis.ul.ie/staff/BF/phd-seminar-series/Hevner-et-al-2004-misq--des-sci.pdf>
38. Hjørland, B. (2008). "Core Classification Theory : a Reply to Szostak." *Journal of Documentation*, **64** (3), 333–342, doi:10.1108/00220410810867560.
39. Hjørland, B. (2010a). "Classification". *Royal School of Library and Information Science*, acessido a 20 de dezembro de 2012, http://www.iva.dk/bh/lifeboat_ko/CONCEPTS/classification.htm.
40. Hjørland, B. (2010b). "Indexing". *Royal School of Library and Informatin Science*, acessido a 7 de janeiro de 2012, http://www.iva.dk/bh/lifeboat_ko/CONCEPTS/indexing.htm.
41. Hjørland, B. (2011). "The importance of theories of knowledge: Browsing as an example". *Journal of the American Society for Information Science and Technology*, **62**(3), 594-603.

42. Hjørland, B. e Pedersen, K. (2005). "A Substantive Theory of Classification for Information Retrieval." *Journal of Documentation* **26** (2), 89–101.
doi:10.1108/00220410510625804.
43. IBM (2012). "FileNet content manager". *IBM Corporation*, acessado a 20 de dezembro de 2012, <http://www-01.ibm.com/software/data/content-management/filenet-content-manager/>.
44. IHMC CmapTools (2013). "Download IHMC CmapTools". Acessado a 25 de março de 2013, <http://cmap.ihmc.us/download/>.
45. Jacob, Elin K. (1991). "Classification and Categorization: Drawing the Line." Em *2nd ASIS SIG/CR Classification Research Workshop*, 63–80.
doi:10.7152/acro.v2i1.12548.
46. Jacob, Elin K. (2001). "The Everyday World of Work: Two Approaches to the Investigation of Classification in Context." *Journal of Documentation*, **57**(1), 76–99.
47. Jacob, Elin K. (2004). "Classification and Categorization: A Difference that Makes a Difference". *Library Trends* **52**(3), 515-540, acessado a 20 de novembro de 2012, <https://www.ideals.illinois.edu/bitstream/handle/2142/1686/Jacob515540.pdf?sequence=2>.
48. Johnston, M. (2012). "People's Choice Winner for Best Open Source CMS is...". *CMSCritic*, acessado a 20 de dezembro de 2012, <http://www.cmscritic.com/peoples-choice-winner-for-best-open-source-cms-is/>.
49. Kampffmeyer, U. (2004). "Trends in Record, Document and Enterprise Content Management". *Project Consult*, Hamburg, acessado a 20 de dezembro de 2012, http://twiki.bes-sa.org/pub/Development/SystemsPlanningInformationSystem/ECM_Handout_english_SER.pdf
50. Kamruzzaman, S. (2010). "Text Classification Using Artificial Intelligence." University of Rajshahi.
51. Kebede, G. (2010). "Knowledge management: an information science perspective". *International Journal of Information Management*, 30, 416-424, acessado a 20 de novembro de 2012, <http://www.sciencedirect.com/science/article/pii/S0268401210000290#>.

52. Kebede, Gashaw (2010). "Knowledge Management: An Information Science Perspective." *International Journal of Information Management* **30**(5), 416–424, doi:10.1016/j.ijinfomgt.2010.02.004.
53. Keshet, Yael (2011). "Classification Systems in the Light of Sociology of Knowledge." *Journal of Documentation* **67**(1), 144–158, doi:10.1108/00220411111105489.
54. Khazraee, Emad e Lin, Xia (2011). "Demystifying Ontology." In *UDC Seminar 2011*, 41–54.
55. Kiu, Ching-Chieh (2011). "TaxoFolk: A hybrid taxonomy–folksonomy structure for knowledge classification and navigation". *Expert Systems with Applications*, 38, 6049-6058, acessado a 25 de outubro de 2012, <http://www.sciencedirect.com/science/article/pii/S0957417410012510>
56. Kwasnik, B.H. (1992). "The role of classification structures in reflecting and building theory". *3rd ASIS SIG/CR Classification Research Workshop*, 63-82, acessado a 20 de novembro de 2012, <https://journals.lib.washington.edu/index.php/acro/article/view/12597>.
57. Kwasnik, B.H. (1999). " The role of classification in knowledge representation and discovery". *Library Trends*, 22-47, acessado a 20 de novembro de 2012, <http://www.asiaa.sinica.edu.tw/~ccchiang/GILIS/LIS/p22-Kwasnik.pdf>.
58. Laserfiche (2012). "Matriz de características". *Laserfiche*, acessado a 20 de dezembro de 2012, <http://www.laserfiche.com/pt-br/Products/FeatureMatrix>.
59. López, Javier (2005). "State of the art: ontologies". Universidad San Carlos de Guatemala, acessado a 25 de março de 2013, http://www.academia.edu/300216/State_of_the_Art_Ontologies
60. Lykke, Marianne; Häj, Anne, e Tudhope, Doug (2011). "Tagging Behaviour with Support from Controlled Vocabulary." In *ISKO UK Conference 2011*.
61. Mai, Jens-Erik (2003). "The Future of General Classification." *Cataloging & Classification Quarterly*, **37**(1-2), 3–12. doi:10.1300/J104v37n01_02.
62. Mai, Jens-Erik (2004a). "Classification in context: relativity, reality, and representation." *Knowledge Organization* **31**(1), 39-44, acessado a 20 de novembro de 2012, http://jenserikmai.info/Papers/2004_ClassificationInContext.pdf.
63. Mai, Jens-Erik (2004b). "Classification of the Web : Challenges and Inquiries." *Knowledge Organization*, **31**(2), acessado a 12 de março de 2012,

http://projects.ischool.washington.edu/chii/Publications/2004_ClassificationOfTheWeb.pdf

64. Mai, Jens-Erik (2010). "Classification in a social world: bias and trust", *Journal of Documentation*, **66**(5), 627 – 642, doi:10.1108/00220411011066763.
65. Mai, Jens-Erik (2011a). "Folksonomies and the New Order : Authority in the Digital Disorder." *Knowledge Organization*, **38**(2): 114–122.
66. Mai, Jens-Erik (2011b). "The Modernity of Classification." *Journal of Documentation*, **67**(4): 710–730, acedido a 27 de março de 2013, http://jenserikmai.info/Papers/2011_Modernity.pdf
67. March, S.T. e Smith, G.F. (1995). "Design and natural science research on information technology". *Decision Support Systems*, 15, 251-266.
68. Matthews, Brian et al. (2009). "An Evaluation of Enhancing Social Tagging with a Knowledge Organization System." Em *ISKO UK Conference 2009*, 22–23.
69. McAfee, a. (2006). "Enterprise 2.0 Inclusionists and Deletionists". *Andrew McAfee's Blog*, acedido a 20 de dezembro de 2012, http://andrewmcafee.org/2006/09/enterprise_20_inclusionists_and_deletionists/.
70. McCool, Matthew (2009). "Field Dependence and Classification: Implications for Global Information Systems." *Journal of The American Society for Information Science and Technology*, **60**(2), 1258-1266.
71. Moore, Andy (2001). "Best Practices in Enterprise Content Management." *KMWorld Special Supplement*, acedido a 20 de dezembro de 2012, http://www.providersedge.com/docs/km_articles/best_practices_in_enterprise_content_management_volume_i.pdf.
72. Munkvold, Bjørn et al. "Contemporary Issues of Enterprise Content Management : The Case of Statoil." Acedido a 20 de dezembro de 2012, <http://is2.lse.ac.uk/asp/aspecis/20030015.pdf>.
73. Novak, Joseph e Cañas, Alberto (2008). "The Theory Underlying Concept Maps and How to Construct and Use Them". Technical Report IHMC Cmap Tools, *Florida Institute for Human and Machine Cognition*, acedido a 25 de março de 2013, <http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf>.
74. Nunes, L. (2007). "Da classificação das ciências à classificação da informação: uma análise do acesso ao conhecimento". Dissertação de mestrado, Pontifícia

Universidade Católica de Campinas, acessado a 20 de novembro de 2012,
<http://alexandria.cpd.ufv.br:8000/teses/ciencia%20da%20informacao/2007/205488f.pdf>.

75. O’Callaghan R., Smits, M. (2005). “A strategy development process for enterprise content management”. Em *Proceedings of the 13th European Conference on information Systems*, Regensburg, Alemanha, 2005, acessado a 2 de janeiro de 2013, <http://is2.lse.ac.uk/asp/aspecis/20050110.pdf>.
76. OpenText (2012). “OpenText ECM Suite”. *OpenText Corporation*, acessado a 20 de dezembro de 2012, <http://www.opentext.com/2/global/products/products-opentext-ecm-suite.htm>.
77. Orna, Elizabeth(2011). “Classification and Visualization of Knowledge . Light from a Forgotten Past.” Em *ISKO UK Conference 2011*.
78. Päivärinta, T., Munkvold, B.E. (2005). “Enterprise Content Management: An Integrated Perspective on Information Management”. Em *Proceedings of the 38th Hawaii International Conference on System Sciences*, EUA, acessado a 25 de outubro de 2012, http://brage.bibsys.no/hia/bitstream/URN:NBN:no-bibsys_brage_2588/1/01385431.pdf.
79. Park, Hee Jin (2009). “Understanding a folksonomy as a web classification.” Tese de doutoramento, acessado a 12 de janeiro de 2013, https://circle.ubc.ca/bitstream/handle/2429/16324/ubc_2010_spring_park_heejin.pdf?sequence=1
80. Park, Heejin (2011). “A conceptual framework to study folksonomic interaction.” *Knowledge Organization*, **38**(6), 515-529.
81. Pinto, M.A., Silva, A.M. (2005). “Um Modelo Sistêmico E Integral De Gestão Da Informação Nas Organizações”. *2º Congresso Internacional de Gestão da Tecnologia e Sistemas de Informação*, São Paulo, Brasil, 2005, acessado a 25 de outubro de 2012, <http://ler.letras.up.pt/uploads/ficheiros/3085.pdf>.
82. Qin, J. (2002). “Evolving Paradigms of Knowledge Representation and Organization: A Comparative Study of Classification, XML/DTD, and Ontology”. Syracuse University, EUA, acessado a 20 de novembro de 2012, http://jqin.mysite.syr.edu/pubs/qin_isko2002.pdf
83. Regli, T; Bloem, A. (2010). “Understanding Content Collection and Indexing”. Information management, acessado a 20 de dezembro de 2012, <http://www.information->

management.com/issues/20_7/content_management_data_integration_indexing_metadata-10019105-1.html.

- 84.** Rouhani, S; Asgari, S.; Mirhosseini, S. (2012). “Review Study: Business Intelligence Concepts and Approaches”. *American Journal of Scientific Research* 50, 62-75, acessado a 20 de dezembro de 2012, http://www.eurojournals.com/AJSR_50_08.pdf.
- 85.** Rudy, M. (2012). “The Top Business Intelligence Software Vendors”. *Compare Business Products*, acessado a 20 de dezembro de 2012, <http://www.comparebusinessproducts.com/business-intelligence/the-top-business-intelligence-software-vendors>.
- 86.** Scott, Judy E (2011). “User Perceptions of an Enterprise Content Management System.” *Em 2011 44th Hawaii International Conference on System Sciences*, 1–9. Ieee. doi:10.1109/HICSS.2011.473.
- 87.** Sem, A. (2004). “Metadata management: past, present and future”. *Decision Support Systems*, 37, 151-173, acessado a 20 de novembro de 2012, <http://el.mdu.edu.tw/datacos/09420322006A/%E8%AB%96%E6%96%87%E9%81%B8%E8%AE%80/Metadata%20management.pdf>.
- 88.** Soergel, Dagobert (1999). “The Rise of Ontologies or the Reinvention of Classification.” *Journal of the American Society for Information Science*, **50**(12).
- 89.** Sowa, John F. (2006). “Semantic networks”. Acessado a 25 de março de 2013, <http://www.jfsowa.com/pubs/semnet.htm>.
- 90.** Srivastava, S. (2012). “What is Business Intelligence?”. *Compare Business Products*, acessado a 20 de dezembro de 2012, <http://www.comparebusinessproducts.com/business-intelligence/what-is-business-intelligence>.
- 91.** Stock, W. G. (2010) “Concepts and semantic relations in Information Science”. *Journal of the American Society for Information Science and Technology*, **61**(10), 1951-1969.
- 92.** Szostak, Rick (2008). “Classification, Interdisciplinarity, and the Study of Science.” *Journal of Documentation*, **64**(3): 319–332. doi:10.1108/00220410810867551.
- 93.** Tanaka, Midori et al. (1995). “Integration of Multiple Knowledge Representation for Classification Problems.” *Artificial Intelligence in Engineering* **9**, 243–251.

94. Tennis, Joseph (2008). "Epistemology, Theory, and Methodology in Knowledge Organization: Toward a Classification, Metatheory, and Research Framework." *Knowledge Organization*, **35**, 102–112.
95. Tennis, Joseph (2011). "Comparative Modeling of Vickery 's Faceted Classification and the Oeuvre of S . R . Ranganathan." In *ISKO UK Conference 2011*.
96. Triska, Ricardo. "Artificial Intelligence, classification theory and the uncertainty reduction process". Federal University of Santa Catarina, acessado a 20 de fevereiro de 2013.
97. T-Systems (2012). "Enterprise Information Management : Get more out of your data". T-Systems, acessado a 20 de dezembro de 2012, <http://www.t-systems.com/solutions/enterprise-information-management-get-more-out-of-your-data/761544>.
98. Tulic, M. (2005). "Automatic indexing". Acessado a 9 de janeiro de 2013, <http://www.anindexer.com/about/auto/autoindex.html>.
99. Vaishnavi, Vijay e Kuechler, Bill (1996). "Design Science Research in Information Systems." *Association for Information Systems*.
100. Van den Heuvel, Charles (2011). "Multidimensional Classifications: Past and Future Conceptualizations and Visualizations ." Em *Proceedings from North American Symposium on Knowledge Organization*, 105–121.
101. Wang, Zhonghong et al. (2008). "Using Classification Schemes and Thesauri to Build an Organizational Taxonomy for Organizing Content and Aiding Navigation." *Journal of Documentation*, **64** (6), 842–876. doi:10.1108/00220410810912424.
102. Weng, Sung-Shun (2006). "Ontology construction for information classification". *Expert Systems with Applications*, **31**, 1-12, acessado a 20 de novembro de 2012, <http://www.sciencedirect.com/science/article/pii/S0957417405001831#>.

Anexos

Anexo 1 - Guião simplificado de aplicação do método para desenvolvimento e gestão de classificações à medida

- 1.** Planeie o processo de desenvolvimento e gestão da classificação
 - 1.1.** Defina a equipa de colaboradores encarregues do desenvolvimento e gestão da classificação
 - 1.2.** Selecione e implemente as plataformas colaborativas para comunicação e discussão entre os membros da equipa (por exemplo, uma wiki ou wiki semântica)
 - 1.3.** Agende as tarefas do processo
- 2.** Defina as necessidades de recuperação de informação dos utilizadores do contexto
 - 2.1.** Analise os conteúdos que pretende classificar quanto à sua natureza (se são para gestão interna do contexto ou se para disponibilizar ao público externo) e quanto à sua utilidade no futuro
 - 2.2.** Definas as necessidades de recuperação de informação do contexto tendo em conta a análise feita aos conteúdos (consulte a lista abaixo para saber quais as necessidades de recuperação correspondentes a cada tipo de conteúdos, e qual a estrutura de classificação mais adequada para cada caso):
 - a.** Conteúdos para gestão interna do contexto – representa a necessidade de recuperação de informação para gestão interna de acordo com a estrutura do contexto (recomenda-se o desenvolvimento de uma estrutura hierárquica de pastas, de acordo com a estrutura do próprio contexto;
 - b.** Conteúdos para disponibilizar ao público externo – representa a necessidade de recuperação de informação a disponibilizar ao público, por utilizadores externos ao contexto (recomenda-se o desenvolvimento de classificação facetada, com a utilização de facetas correspondentes às possíveis necessidades do público-alvo);
 - c.** Conteúdos com utilidade no futuro – *representa a necessidade de recuperação de informação no âmbito da gestão do conhecimento para o futuro* (recomenda-se o desenvolvimento de uma classificação facetada, com a utilização de facetas correspondentes às necessidades dos colaboradores do contexto (por exemplo, “área”, “serviço”, “tipo de documento”, “versão do documento”, entre muitas outras));
 - d.** Necessidade extra: necessidade de recuperação de informação de acordo com necessidades pessoais dos colaboradores e manutenção de uma rede social de partilha de informação (recomenda-se o sistema de tagging, onde o utilizador pode atribuir tags aos conteúdos de acordo com o que pretender);

3. Desenvolva a classificação segundo uma das seguintes abordagens: Abordagem genérica é recomendada caso o contexto seja de pequena dimensão, todos os colaboradores possam integrar a equipa e contribuir no processo e não existam muitas restrições de tempo para concluir o processo; Abordagem especializada é recomendada caso o contexto seja de grande dimensão, a equipa integre apenas alguns colaboradores, e haja restrições de tempo mais rigorosas.

Abordagem genérica

- 3.1.** Crie um mapa de conceitos genérico do domínio do contexto, através da combinação de vários mapas de conceitos individuais criados por cada colaborador (é importante que cada colaborador tente representar todas as perspetivas possíveis do domínio);
- 3.2.** Defina a estrutura para a classificação de acordo com as necessidades de recuperação definidas (consulte recomendações do ponto 2.2.)
- 3.3.** Utilize partes do mapa de conceitos genérico para definir os termos da classificação e a sua organização na classificação

Abordagem especializada

- 3.1.** Defina a estrutura para a classificação de acordo com as necessidades de recuperação definidas (consulte recomendações do ponto 2.2.)
 - 3.2.** Defina os termos e a sua organização na classificação (pode-se basear em tesauros e em classificações já existentes para o domínio)
- 4.** Avalie a classificação
- 4.1.** Teste a classificação numa amostra de conteúdo, verificando se a classificação:
 - a.** Tem termos suficientes
 - b.** É de fácil utilização
 - c.** Tem potencial para ser útil para a gestão da informação no contexto
 - 4.2.** Corrija e reformule a classificação, se necessário, até cumprir os três requisitos do ponto 4.1
- 5.** Implemente e classificação
- 5.1.** Adapte a classificação às plataformas de gestão de conteúdos (utilize pastas para adaptar classificações em hierarquia, e etiquetas/tags para adaptar classificação facetadas);
 - 5.2.** Crie um guia de utilização da classificação para os colaboradores
 - 5.3.** Classifique o conteúdo já existente na(s) plataforma(s) de gestão de conteúdos
- 6.** Faça a gestão e atualização da classificação

- 6.1. Monitorize a utilização da classificação pelos colaboradores através da análise do feedback dos colaboradores, fornecido se possível através das plataformas colaborativas
- 6.2. Atualize a classificação conforme necessário e corrija os erros/problemas encontrados pelos colaboradores

Anexo 2 - Versão provisória e resumida do método, utilizada na avaliação

FASE	ETAPA	DESCRIÇÃO
PLANEAMENTO	Preparação	✓ Definição dos recursos ○ Definição da equipa encarregue do desenvolvimento e gestão da classificação ○ Seleção (e implementação) do sistema de gestão de conteúdos e das plataformas eletrónicas que farão uso da classificação ○ Seleção (e implementação) das plataformas colaborativas (meio de comunicação e colaboração entre os membros da equipa) ✓ Estabelecimento dos objetivos e potenciais benefícios da classificação para a gestão da informação ✓ Análise do tipo de conteúdos a ser integrado no sistema de conteúdos ○ É informação para gestão interna, ou informação a disponibilizar ao público, ou informação registada para gestão do conhecimento e utilidade noutros contextos futuros? ✓ Gestão do processo de desenvolvimento da classificação ○ Calendarização de cada etapa do método ○ Realização de tarefa no sentido de motivar/incentivar a colaboração da equipa (tais como comunicação através das plataformas colaborativas (p. ex. wikis), comunicação dos benefícios da classificação, registo e discussão dos problemas/dúvidas, etc.) Resultados pretendidos nesta etapa: • Equipa definida, sistema de gestão de conteúdos e outras plataformas implementadas, e plataformas colaborativas implementadas • Objetivos e benefícios da classificação definidos e comunicados à equipa • Definição do tipo de conteúdos a ser integrado no sistema • Calendarização das etapas do método definidas e incentivo contínuo à colaboração da equipa
	Definição das necessidades de recuperação de informação	✓ Definição das necessidades de recuperação de informação ○ <u>Recuperação de informação “interna” de acordo com a estrutura do contexto</u> – recomenda-se o modelo estrutural de um sistema hierárquico por pastas, de acordo com as atividades, processos ou tarefas realizada no contexto. (Se a quantidade de conteúdos a gerir se tornar muito elevada (superior a 1000 objetos informacionais), o tempo de recuperação poderá começar a tornar-se bastante longo, e começará a ser necessária a implementação de um modelo estrutural mais completo, como a classificação facetada) ○ <u>Recuperação de informação “pública”</u> – recomenda-se o modelo estrutural de classificação facetada, com a utilização das facetas necessárias para corresponder às possíveis necessidades do típico utilizador de conteúdos “públicos” ○ <u>Recuperação de informação no âmbito da gestão do conhecimento para o futuro</u> – recomenda-se o modelo estrutural de classificação facetada, com a utilização das facetas necessárias para corresponder às possíveis necessidades do utilizador-colaborador do contexto ○ <u>Recuperação de informação de acordo com necessidades pessoais e dando incentivo à colaboração e participação na partilha de conhecimento</u> – recomenda-se o modelo estrutural da folksonomia (sistema de tagging), permitindo que o utilizador atribua tags ao conteúdo de acordo com a sua perspetiva pessoal da informação Resultados pretendidos nesta etapa: • Necessidades de recuperação de informação encontradas
DESENVOLVIMENTO	Desenvolvimento da classificação (optar por uma das abordagens – a abordagem	<u>Abordagem especializada</u> <u>Abordagem genérica</u> 1. Construção de rede semântica genérica 1.1. Construção de um mapa de conceitos relacionados entre si, e

	especializada é mais adequada para satisfazer necessidades de recuperação específicas; a abordagem genérica é mais adequada quando se pretende criar uma base de conhecimento do qual se poderão criar classificações de acordo com as necessidades presentes ou futuras)	1. Definição do modelo estrutural 1.1. Escolha do modelo estrutural de classificação mais adequado para satisfazer as necessidades de recuperação de informação definidas na etapa anterior (consulte a informação da etapa anterior) 2. Construção da classificação especializada 2.1. Definição dos conceitos relevantes para a classificação especializada, e escolha dos termos (recomenda-se a consulta de tesouros e de outras classificações relacionadas com a área do contexto. Também pode ser útil a construção de um mapa de conceitos para organizar o conhecimento) 2.2. Organização dos termos através da definição das relações (na classificação facetada esta fase corresponde à criação de facetas (“tipo de documento”, “processo”, “atividade”, “assunto”, etc.)) 2.3. Atribuição de notações a cada classe (opcional) Resultados pretendidos nesta etapa: • Classificação estruturada e organizada de acordo com as necessidades de recuperação encontradas na segunda etapa, e pronta para ser testada através da aplicação da mesma a uma amostra de conteúdo	relevantes para o contexto. Devem ser consideradas todas as perspetivas possíveis do contexto, de forma a se obter uma rede genérica. As típicas relações entre conceitos serão “x é um y” e “x tem y”. A colaboração e contribuição individual dos membros da equipa são vitais nesta tarefa (recomenda-se a utilização das aplicação CmapTools ou ConceptME) 2. Definição dos modelos estruturais 2.1. Escolha dos modelos estruturais de classificação mais adequados para satisfazer as necessidades de recuperação de informação definidas na etapa anterior (consulte a informação da etapa anterior) 3. Derivação da rede semântica para classificação especializada 3.1. Adaptação de parte da rede semântica genérica para a criação de uma classificação especializada utilizando o modelo estrutural adequado à necessidade de recuperação de informação existente no contexto (ou seja, utilizar uma parte dos conceitos do mapa de conceitos para criar uma classificação especializada, e cujo modelo estrutural deverá ser selecionado de acordo com as necessidades de recuperação encontradas na segunda etapa. Normalmente, as relações definidas como “x é um y” resultam em classificações hierárquicas, e as relações “x tem y” resultam em classificação facetadas, onde “y” é uma faceta de “x”) 3.2. Atribuição de notações a cada classe (opcional)
Avaliação	1. Teste da classificação através da aplicação da mesma a uma amostra de conteúdo (classificar, em forma de teste, uma amostra do conteúdo típico a ser gerido no sistema) 1.1. Avaliação da funcionalidade (termos suficientes/insuficientes), usabilidade (facilidade ou dificuldade de aplicação da classificação) e utilidade (se é provável/improvável) que a classificação venha melhorar a recuperação dos conteúdos) da classificação 1.2. Identificação e registo dos problemas encontrados		

		2. Possível correção/reformulação da classificação 2.1. Correção dos problemas da classificação – fraca usabilidade é sintoma de elevada complexidade ou profundidade na classificação; pouca utilidade é sintoma de erros na definição das necessidades de recuperação de informação ou na escolha do modelo estrutural mais adequado; fraca funcionalidade (termos insuficientes para classificar todo o tipo de conteúdos) é sintoma de pouca profundidade na classificação e de esforço insuficiente na definição dos conceitos e relações relevantes 2.2. Reavaliação da classificação Resultados pretendidos nesta etapa: Classificação avaliada de forma positiva e pronta para ser adaptada e implementada no sistema de gestão de conteúdos
	Implementação	✓ Adaptação da classificação às funcionalidades do sistema de conteúdos - O modelo estrutural de classificação pode ter que ser “modificado” quando aplicado às funcionalidades de classificação do sistema de conteúdos. Os sistemas de conteúdos não são, na maioria das vezes, capazes de suportar classificação mais complexas, como as facetadas. Por isso, recomenda-se que se mantenha a classificação estabilizada externamente ao sistema (por exemplo, na plataforma colaborativa utilizada para gerir o processo de desenvolvimento da classificação), e que se faça a adaptação necessária aos sistemas onde a classificação será implementada. Normalmente, a adaptação passará pela criação de uma estrutura hierárquica de pastas (num modelo estrutural em hierarquia), ou pela atribuição de categorias a cada conteúdo de acordo com as várias facetas da classificação (no caso de uma classificação facetada, a faceta “principal” pode ser transposta para uma estrutura de pastas, e as restantes serem adaptadas através da atribuição de categorias). No caso das folksonomias, as tags serão também atribuídas através das opções de atribuição de categorias. No entanto, caso os sistemas utilizados disponham de funcionalidades mais complexas que a simples atribuição de categorias ao conteúdo, poderão surgir novas formas de levar a cabo esta adaptação. ✓ Modelação do processo de classificação do conteúdo e comunicação do mesmo aos colaboradores do contexto (criação de um guião sobre como classificar o conteúdo no sistema) ✓ Classificação do conteúdo existente no sistema segundo a classificação desenvolvida Resultados pretendidos nesta etapa: - Classificação adaptada e implementada no sistema de gestão de conteúdos - Criação e disponibilização aos colaboradores de um guião sobre como classificar o conteúdo no sistema - Conteúdo armazenado no sistema classificado de acordo com a classificação desenvolvida
GESTÃO	Gestão e atualização	✓ Análise do feedback dos colaboradores acerca da utilização da classificação (através da utilização das plataformas colaborativas) - Monitorização da utilidade da classificação (se satisfaz/não satisfaz as necessidades de recuperação encontradas na segunda etapa) - Monitorização de possíveis novas necessidades de recuperação de informação - Monitorização de incapacidades da classificação (termos insuficientes) - Monitorização de erros na semântica da classificação (termos ou relações mal enquadrados) ✓ Atualização da classificação - Correção de erros na classificação atribuída aos conteúdos - Inclusão de novos termos e relações - Necessidade de novo modelo estrutural para satisfazer necessidade de recuperação de informação emergente

Anexo 3 - Formulário para o registo de informação na avaliação

Informação a registar durante a aplicação do método:

- Problemas encontrados e dificuldades experienciadas durante a realização de cada tarefa
- Desvios feitos em relação às indicações do método e razões/causas para os mesmos (o modelo não impõe uma aplicação rigorosa, os desvios são possíveis e até necessários de acordo com cada contexto)
- Resultados obtidos em cada etapa
- Outras informações que pareçam relevantes para a avaliação do método

ETAPA	TAREFAS E RESULTADOS	REGISTO DO PROCESSO
	✓ Definição dos recursos	Recursos: Serviço de consultoria
	✓ Estabelecimento dos objetivos e potenciais benefícios da classificação para a gestão da informação	Informação para gestão interna (durante a execução do projeto), informação registada para gestão do conhecimento e utilidade noutros contextos futuros (durante e pós projeto).
	✓ Análise do tipo de conteúdos a ser integrado no sistema de conteúdos	Tipo de conteúdos: - conteúdos técnicos - conteúdos de gestão de projeto
	✓ Gestão do processo de desenvolvimento da classificação	✓ Problemas/dificuldades - ✓ Desvios - ✓ Outras informações -
	<u>Resultados pretendidos nesta etapa:</u> • Equipa definida, sistema de gestão de conteúdos e outras plataformas implementadas, e plataformas colaborativas implementadas • Objetivos e benefícios da classificação	✓ Resultados obtidos:

	definidos e comunicados à equipa - Definição do tipo de conteúdos a ser integrado no sistema - Calendarização das etapas do método definidas e incentivo contínuo à colaboração da equipa		
Definição das necessidades de recuperação de informação	✓ Definição das necessidades de recuperação de informação - Recuperação de informação “interna” de acordo com a estrutura do contexto - Recuperação de informação “pública” - Recuperação de informação no âmbito da gestão do conhecimento para o futuro - Recuperação de informação de acordo com necessidades pessoais e dando incentivo à colaboração e participação na partilha de conhecimento		
	Resultados pretendidos nesta etapa: Necessidades de recuperação de informação encontradas		Resultados obtidos (quais as necessidades encontradas)? - Recuperação de informação “interna” de acordo com a estrutura do contexto –o modelo estrutural de um sistema hierárquico por pastas, de acordo com as tarefas e resultados do projeto, - Recuperação de informação no âmbito da gestão do conhecimento para o futuro – utilização de tags?...
Desenvolvimento da classificação (optar por uma das abordagens – a abordagem especializada é mais adequada para satisfazer necessidades de recuperação específicas; a abordagem genérica é mais adequada quando se pretende criar uma base de conhecimento do qual se poderão criar classificações de acordo com as necessidades presentes ou futuras)	Abordagem especializada	Abordagem genérica	Qual a abordagem escolhida? Existiram dificuldades na escolha? Abordagem especializada. A área de Consultoria já dispõe de uma estrutura de pastas standard (<i>ver anexo</i>), utilizada para todos os projetos de consultoria, orientada aos resultados e às tarefas a executar para atingir esses resultados. Cada documento apresenta também uma designação standard, nomeadamente: [SiglaProjeto].tp.ff.pp.parceiro.aaaammdd.ddddd.VD.d Os seus elementos permite obter informação sobre: Tp: tipo de documento Ff: fase do projeto Pp: processo Parceiro: ex. cliente ou INESC Aaaammdd: data VD.d: versão. Para a Gestão do Conhecimento, foi criada anteriormente uma classificação do tipo facetada para permitir a pesquisa de informação numa ótica mais a longo prazo. A classificação foi elaborada num trabalho

		1. Construção de rede semântica genérica	de dissertação da Diana Vieira no ano passado (<u>ver anexo VII da dissertação que te envio</u>).
			✓ Problemas/dificuldades (se escolher abor. Genérica) - ✓ Desvios (se escolher abor. Genérica) - ✓ Outras informações (se escolher abor. Genérica) -
	1. Definição do modelo estrutural	2. Definição dos modelos estruturais	✓ Problemas/dificuldades - ✓ Desvios - ✓ Outras informações -

	2. Construção da classificação especializada	3. Derivação da rede semântica para classificação especializada	✓ Problemas/dificuldades - ✓ Desvios - ✓ Outras informações -
	Resultados pretendidos nesta etapa: Classificação estruturada e organizada de acordo com as necessidades de recuperação encontradas na segunda etapa, e pronta para ser testada através da aplicação da mesma a uma amostra de conteúdo		✓ Resultados obtidos (qual o modelo estrutural escolhido, que tipo de conceitos e termos inclui, etc.)
Avaliação	1. Teste da classificação através da aplicação da mesma a uma amostra de conteúdo (classificar, em forma de teste, uma amostra do conteúdo típico a ser gerido no sistema) - Avaliação da funcionalidade, usabilidade e utilidade (da classificação) - Identificação e registo dos problemas encontrados		✓ Problemas/dificuldades - ✓ Desvios - ✓ Outras informações -

	2. Possível correção/reformulação da classificação a. Correção dos problemas da classificação b. Reavaliação da classificação	✓ Problemas/dificuldades - ✓ Desvios - ✓ Outras informações -
	Resultados pretendidos nesta etapa: Classificação avaliada de forma positiva e pronta para ser adaptada e implementada no sistema de gestão de conteúdos	✓ Resultados obtidos (qual os aspetos positivos e negativos encontrados na classificação desenvolvida, e se a mesma é capaz de corresponder aos objetivos)
Implementação	✓ Adaptação da classificação às funcionalidades do sistema de conteúdos	✓ Problemas/dificuldades - ✓ Desvios - ✓ Outras informações -
	✓ Modelação do processo de classificação do conteúdo e comunicação do mesmo aos colaboradores do contexto	✓ Problemas/dificuldades - ✓ Desvios -

		✓ Outras informações -
	✓ Classificação do conteúdo existente no sistema segundo a classificação desenvolvida <u>Resultados pretendidos nesta etapa:</u> • Classificação adaptada e implementada no sistema de gestão de conteúdos • Criação e disponibilização aos colaboradores de um guião sobre como classificar o conteúdo no sistema • Conteúdo armazenado no sistema classificado de acordo com a classificação desenvolvida	✓ Problemas/dificuldades - ✓ Desvios - ✓ Outras informações - ✓ Resultados obtidos (descrição breve da forma como a classificação foi/será adaptada ao sistema)
Gestão e atualização	✓ Análise do feedback dos colaboradores acerca da utilização da classificação (através da utilização das plataformas colaborativas) ○ Monitorização da utilidade da classificação (se satisfaz/não satisfaz as necessidades de recuperação encontradas na segunda etapa) ○ Monitorização de possíveis novas necessidades de recuperação de informação ○ Monitorização de incapacidades da classificação (termos insuficientes) ○ Monitorização de erros na semântica da classificação (termos ou relações mal enquadrados) ✓ Atualização da classificação ○ Correção de erros na classificação atribuída aos conteúdos ○ Inclusão de novos termos e relações ○ Necessidade de novo modelo estrutural para satisfazer necessidade de recuperação de informação emergente	(esta fase não poderá ser avaliada, uma vez que requer um período de tempo bastante grande para se poderem registar mudanças no contexto organizacional que proporcionem alterações na classificação)

Anexo 4 - Transcrição das entrevistas realizadas

Entrevista a “anónimo 1”

[...]

Qual o processo normal no momento de “armazenamento/submissão” de um documento no vosso sistema de gestão de conteúdos?

Anónimo 1: Portanto, toda a documentação que nós criamos no âmbito do projeto...nós temos documentação técnica nossa e que vai acompanhando o projeto e depois também existe uma parte da documentação que é... alguns deliverables. Nós fazemos isso através de uma plataforma, que por vezes é um repositório de ficheiros, que está organizado por pastas correspondentes às tarefas ou aos resultados de projetos, e os documentos vão sendo inseridos nessas pastas. Não estamos a catalogar, não estamos a identificar..

Só colocam na pasta respetiva...

Anónimo 1: Exato. Há umas pastas que, durante o projeto, identificam as grandes atividades, e depois dentro destas podem ainda existir subpastas quando existem tarefas. Dentro destas, cada colaborador está livre para criar as pastas que quiser, apesar de existir sempre um acompanhamento da organização que é feita. E nestas estão os documentos que nós vamos produzindo ao longo dos projetos. Noutra pasta à parte estão os resultados obtidos durante o projeto, que são enviados às direções, e outros resultados que possam surgir, como software, etc. A estrutura também existe na Dropbox, no entanto a Dropbox serve apenas para partilha interna de ficheiros... continua a ser necessário existir um repositório organizado para a informação geral do projeto..

E quanto à parte da documentação técnica que falou?

Anónimo 1: Os outros documentos também ficam armazenados no sistema de pastas...

O mesmo que falou anteriormente?

Anónimo 1: Exatamente.

Não existem então duas estruturas hierárquicas? É tudo armazenado na mesma estrutura?

Anónimo 1: Sim

Bem, a pergunta que se impõe a seguir é: quais os problemas que têm encontrado com a estruturação que foi feita?

Anónimo 1: Recentemente, estamos a tentar implementar outra forma. Estamos a utilizar uma plataforma que permite – e é também um repositório de ficheiros – que toda a gente possa fazer o controlo de versões dos documentos, e onde também é possível ir mantendo um wiki com informação ou conhecimento que...

Está a falar do *TWiki*?

Anónimo 1: Não, mas é equivalente. Faz basicamente a mesma coisa...

E a outra plataforma que normalmente era utilizada é o *Plone*? Penso que é o que é mais comum nos projetos do INESC...

Anónimo 1: Sim. Mas mesmo o *Plone*, que já utilizei num projeto, foi também apenas utilizado como repositório de ficheiros, mais nada. No entanto o *Plone* permitia-nos ter uma área para gestão interna do projeto, e uma área externa pública. Os colaboradores, contudo, acharam complicado trabalhar com o *Plone* e outras plataformas semelhantes, principalmente pela dificuldade em perceber onde é que os documentos estão... e depois tinham de fazer download dos ficheiros para edição local, e voltar a fazer o *upload* para a plataforma. Nos projetos onde utilizamos essa plataforma, verificamos que as empresas preferiam receber o ficheiro por e-mail, fazer as alterações, devolver-nos novamente por e-mail, e nós é que fazíamos o *upload*...

[...] normalmente a submissão de ficheiros era normalmente feita por uma única pessoa, enquanto as trocas durante a elaboração do documento era normalmente feita através de outro sistema, como a Dropbox, ou o e-mail...

Anónimo 1: É isso... eu penso que os sistemas não são muito amigáveis. Um exemplo é o facto de não deixarem fazer a edição online dos documentos – se não fosse necessário descarregar, editar e voltar a submeter...

Pois... a edição online só é possível quando os documentos são criados diretamente na plataforma...

Anónimo 1: Sim, mas isso torna-se complicado por causa dos *templates word* que são criados para normalizar os documentos do projeto... perder-se-ia muito tempo na edição. Pronto, agora nós temos então o Wiki que está associado a um projeto e onde as pessoas devem ir colocando alguma informação, algumas notas do que estão a fazer, e alguns detalhes técnicos do trabalho que vão desenvolvendo...

E essa plataforma já começa a ser mais utilizada pelos colaboradores?

Anónimo 1: Estamos a tentar agora criar algumas regras e impor que as pessoas utilizem. Nós sentimos que estas equipas mudam com muita facilidade, uma pessoa pode estar a trabalhar num projeto, e de repente já “saltar” para outro... e perde-se todo o conhecimento que essa pessoa tinha adquirido.

Pois... e normalmente as wikis são utilizadas muito de forma ad-hoc, e não existe a preocupação em escrever a informação de forma a esta poder ser lida e interpretada corretamente por outro colaborador no futuro

Anónimo 1: Exato... não há aquela disciplina de ir guardando o que se está a fazer, o que se consultou, o processo que realizou, e o que decidiu. Por vezes existe é uma imposição do responsável do projeto, no final, para registar determinada parte, e aí realmente a pessoa senta-se e começa a escrever qualquer coisa... claro que já não é feito com o detalhe que seria feito se fosse durante a realização da própria tarefa... mas sempre se resume alguma coisa.

Então no *Plone* e nas plataformas que normalmente utilizavam nem era sequer feito o registo de qualquer conhecimento por parte dos colaboradores?

Anónimo 1: Que eu tenha conhecimento não. Eu sinto que o facto de toda a gente estar a trabalhar no mesmo sítio em equipa, faz com que se registe menos as coisas. Se as pessoas trabalharem de forma mais dispersa, existe uma maior preocupação com o registo do que vai sendo feito e das dificuldades...

Pois, até pela necessidade de comunicar à distância... faz com que as pessoas precisem de escrever para explicarem melhor o que vão fazendo...

Anónimo 1: Quanto estão em equipa no mesmo sítio, temos de ser nós a pedir para registarem qualquer coisa no final do dia..

Isso até estará mais relacionado com a própria cultura do colaborador... antigamente trabalhavam de uma forma e agora vêm-se obrigados a estar presentes no wiki e etc. Penso que no futuro talvez o colaborador esteja mais motivado a ter um tipo de comportamento mais ativo no registo do conhecimento, muito pelo facto de os jovens adultos passarem muito tempo a partilhar informação e a discutir em fóruns e redes sociais...

Anónimo 1: Pois. Nós estamos a tentar mudar essa cultura e essa forma de trabalhar...mas não é fácil.

Pois... isso afeta muito o conhecimento que é transferido de projeto para projeto, o que está muito relacionado com a forma como o conhecimento foi gerido durante o desenvolvimento do projeto. Em termos de problemas de recuperação de informação, têm verificado alguns durante os projetos em que esteve envolvido? Segundo o César, a quantidade de documentos reduzida fazia com que fosse sempre relativamente fácil encontrar determinado documento, no entanto penso que uma maior quantidade de documentos (como numa organização) faria com que essa estrutura hierárquica e o facto de não atribuírem meta-informação aos documentos fariam com que se verificassem muitos problemas de recuperação.

Anónimo 1: Eu também penso isso. O facto do número de documentos que criámos ser reduzido e a forma como a estrutura de pastas está organizada (atividades, resultados do projeto), faz com que a recuperação dos documentos seja simples e sem problemas. No entanto, se pensássemos que criar um sistema onde colocássemos toda a informação dos projetos junta, seria necessário fazer essa classificação dos conteúdos.

A estrutura de pastas foi criada e discutida colaborativamente em equipa, ou a sua organização foi imposta pela direção do projeto?

Anónimo 1: Normalmente é decidida pela direção.

E em termos de reuniões, a gestão de conteúdos era tema de discussão?

Anónimo 1: Sim, mas essa gestão vai sendo estabilizada. Nós temos um procedimento que, ao iniciar um novo projeto, são criados automaticamente a estrutura de pastas (nível mais alto) e os *templates* para os documentos. Isso já é do conhecimento de todas as pessoas que estão a trabalhar no projeto, e portanto não existe discussão quanto a esses temas. Houve no passado, mas neste momento já toda a gente sabe que a estrutura está estabilizada.

E em termo de notação de cada pasta?

Anónimo 1: Em termos de código não, mas as pastas têm sempre uma designação típica. Por vezes, também atribuído um certo código aos resultados –R1, R2. Além disso, existem também um código específico para as atas e convocatórias, que é normalmente referido no assunto dos e-mails para ser de identificação mais fácil. E depois existem sempre os templates típicos dos documentos para o colaborador perceber exatamente do que se trata (se é resultado,..)

Esses *templates* é uma tentativa de normalizar os documentos que são criados no âmbito do projeto?

Anónimo 1: Sim, existem os *templates* específicos, cada projeto tem uma série deles. Tem sempre o logo do parceiro, o nome, etc. Quando os documentos são criados por vários parceiros, existe um *template* específico.

Como fazem o incentivo à utilização das plataformas colaborativas, como o wiki que falou?

Anónimo 1: Normalmente são feitas em reuniões, e principalmente em conversas quando estamos a trabalhar em equipa no mesmo sítio. Há o incentivo a registar no wiki resumidamente o que fizeram, especialmente quando se percebe que existe informação crítica que deve ser registada.

E há melhoria nesse aspeto colaborativo?

Anónimo 1: Ainda não muito...

A informação dos e-mails é gerida de forma central?

Anónimo 1: Não, cada pessoa gere pessoalmente o seu e-mail. Apesar de existirem regras no envio de e-mail.

[...]

Entrevista a “anónimo 2”

[...]

Qual é o processo normal no momento de armazenar ou submeter um conteúdo no sistema de gestão de conteúdos? Penso que a plataforma típica utilizada aqui no INESC é o Plone...

Anónimo 2: Isso pode divergir de projeto para projeto... mas pode explicar melhor a pergunta?

O que pretendo saber é se existe alguma preocupação [da parte do utilizador] quando se coloca um conteúdo no sistema, seja em relação a títulos normalizados, seja...

Anónimo 2: Sim, há preocupações com o título... e coloca-se cada documento numa determinada pasta que já identifica de certa forma o contexto do projeto.

O sistema de pastas já está organizado à partida ou é construído [durante o projeto]?

Anónimo 2: Esse sistema de pastas é específico para cada projeto, embora se possa pensar numa estrutura global para todos os projetos. Mas a estrutura hierárquica de pastas específica para o projeto é normalmente feita pela entidade que organiza o sistema de conteúdos.

E essa entidade inclui pessoal específico envolvido no projeto?

Anónimo 2: No contexto dos projetos europeus, onde eu normalmente atuo, cada projeto define o seu sistema de gestão de conteúdos, define a sua organização e depois um dos parceiros do projeto fica encarregue da implementação do sistema.

A organização do sistema, ou repositório, fica mais na responsabilidade do pessoal informático envolvido no projeto, então?

Anónimo 2: Não, a direção do projeto define um núcleo de parceiros para o projeto, três ou quatro normalmente, que ficam encarregues das decisões acerca da organização de conteúdos do sistema, ou seja, qual informação que deve ser armazenada nesse sistema. Qual a informação pública, qual a informação privada, etc. Depois, alguém ficará encarregue dessa implementação e, por vez, nem sequer é alguém ligado diretamente ao projeto.

Mas, [voltando à questão dos títulos], todos os colaboradores envolvidos no projeto devem estar a par dessa normalização, para além da própria organização da estrutura de pastas?

Anónimo 2: Sim, cada colaborador deve saber em qual pasta da hierarquia deve colocar cada documento.

E o título de cada documento deve ser similar ao próprio nome da pasta, certo? Ou seja, não há probabilidade de haver, por exemplo, dois conteúdos iguais submetidos por dois colaboradores com nomes diferentes?

Anónimo 2: Eu vou mostrar um exemplo... [utilizando o computador] normalmente estes projetos europeus estão organizados por work packages, e depois, por exemplo, aqui na gestão de projetos encontramos todos os documentos que foram criados no âmbito da gestão de projetos. Aqui na Work Package 3, podemos encontrar alguns dos *deliverables* previstos no projeto, algumas minutas sobre workshops técnicos, algumas versões draft dos *deliverables*, e outros documentos que foram utilizados para a construção do draft. E não se dá normalmente o caso de dois documentos idênticos com nomes diferentes e armazenados em pastas diferentes...

Sim, parece que a notação que é utilizada [apontado para uma das pastas] também ajuda...

Anónimo 2: Sim, e o título também é semelhante à designação da pasta.

O que parece acontecer é que as opções de atribuição de categorias disponíveis pelo Plone não são utilizadas para classificar os conteúdos. Na análise que fiz a dois portais utilizados em dois projetos [em que o INESC Porto esteve envolvido] reparei que

apesar de terem sido criadas categorias pré-definidas para os conteúdos, não estavam a ser associadas pelos colaboradores do projeto aos conteúdos que armazenavam.

Anónimo 2: Pois não...

E isso não traz problemas de recuperação de informação?

Anónimo 2: Normalmente quem está envolvido no projeto sabe exatamente em que parte da estrutura de pastas está o documento que precisa, [e chega até ele através de uma navegação pela estrutura]. Como cada colaborador sabe exatamente em que work package está exatamente a trabalhar, não se verificam problemas seja na submissão de um conteúdo ou na recuperação de informação.

Está organizado de forma muito lógica. Provavelmente o facto de o sistema não conter uma elevada quantidade de conteúdos também contribui para que não hajam problemas.

Anónimo 2: Sim, também. Nos projetos europeus, a estrutura de pastas é suficiente [para as nossas necessidades de recuperação de informação], desde que esteja organizada de acordo com o contexto do projeto. O que implica que a estrutura seja dinâmica de forma a incluir pastas relativas a tarefas que não estavam previstas no início do projeto. O problema do dinamismo é o facto de nem todos os colaboradores terem permissões para criarem novas pastas, o que implica que tenhamos que estar a pedir à entidade...

É um dos problemas de subcontratar a implementação do sistema.

Anónimo 2: Pois, constitui um problema.

Parece-me que a gestão de informação neste tipo de projetos é feita de uma forma relativamente ad hoc, ou seja, quando é preciso é necessário estar a pedir...

Anónimo 2: Sim, sim. Mas voltando à questão inicial, a questão de não ser feita a classificação dos conteúdos nunca constituiu grandes problemas [de recuperação de informação] para mim.

Entendo, mas a própria estrutura hierárquica de pastas pode ser considerada uma classificação.

Anónimo 2: Sim, e neste contexto acaba por ser bastante eficiente.

E o próprio facto de não serem armazenados muitos conteúdos no sistema também ajuda.

Anónimo 2: Claro, nunca são muitos documentos. Não me recordo de nenhuma situação onde tivéssemos mais de 1000 documentos no âmbito de um projeto.

Esses conteúdos são elaborados de forma colaborativa e geridos e partilhados entre colaboradores através do próprio sistema de gestão de conteúdos?

Anónimo 2: Não isso depende dos projetos. Antigamente, utilizávamos sempre o correio eletrónico na partilha de documentos. Quando se começaram a utilizar estas plataformas tipo Plone, houve tentativas de utilizar este sistema como forma de elaborar de conteúdos de forma colaborativa, mas não correu bem. Atualmente, utilizamos este novo tipo de ferramentas como a Dropbox. [No fundo], o Plone acaba por ser utilizado apenas como repositório das versões finais dos conteúdos, uma vez que a Dropbox acaba por ser muito mais prática. Este tipo de sistemas como Plone acaba por ser muito “pesado”, e como uma interface nada prática, que obriga a muito “scroll”.

Já se verificou algum problema grave na recuperação de algum conteúdo?

Anónimo 2: Não, contudo por vezes existem alguns problemas na gestão das versões dos documentos. Existem situações em que não sabemos onde está a versão correta dos documentos que precisamos para trabalhar, porque os colaboradores não fazem a atualização do documento no sistema...

A submissão dos documentos no sistema Plone é feita livremente por cada colaborador?

Anónimo 2: Normalmente cada work package tem um líder. Se a determinada altura, numa reunião, se definir que é necessário atualizar algum aspeto, o próprio líder trata dessa situação ou então alguém que ele nomeia.

A gestão da informação no projeto é tema regular nas reuniões?

Anónimo 2: De vem em quando...

Quando há problemas apenas?

Anónimo 2: E não só. Por vez há necessidade de criar pastas novas [para determinada tarefa] e esse tipo de decisões são feitas numa reunião. Este tipo de sistemas tipo Plone tem algumas vantagens em relação à Dropbox, nomeadamente o envio de e-mails a avisar que existiu uma nova submissão ou alteração, e o facto de possuir no mesmo portal uma zona privada, e uma zona pública. Ou seja, é mais conveniente em certas coisas em relação à Dropbox, mas não é tão prático noutras.

[...]

Entrevista a “anónimo 3”

[...]

Anónimo 3: [Os colaboradores da equipa do projeto] estão a usar [o conceptME] na área da Gestão Florestal, e tinham um site construído numa wiki semântica da MediaWiki para organizar tudo sobre sistemas de apoio à decisão dessa área. A wiki semântica é uma plataforma que permite que eles definam uma classificação para classificar as páginas da wiki. Como eles não estavam satisfeitos com a organização que tinham, nós tentamos utilizar a ferramenta ConceptME com eles para os ajudar. A grande diferença entre o ConceptME e, por exemplo, o CmapTools, é que a primeira permite que duas equipas trabalhem primeiro individualmente nas suas ideias e no seu mapa de conceitos, e depois a ferramenta permite agregar os dois mapas. O mapa de conceitos pode depois ser utilizado para criar uma classificação para gerir um sistema de gestão de conteúdos. O conceptME não tenta apenas ser uma ferramenta de automatização, mas sim algo que incentive à discussão. Isto é importante porque mesmo dois especialistas da área podem ter duas visões diferentes do conhecimento do domínio, e assim a ferramenta pretende também apoiar essa discussão.

[...]

Anónimo 3: De certa forma, a maioria dos projetos acaba por ter uma estrutura mais ou menos normalizada, que normalmente está organizada de acordo com a proposta de projeto que nos chega. As work packages do projeto são normalmente sequenciais, e estão normalmente estruturadas de forma a estarem contidas e não terem muita interação com outras work packages. Esta forma de fazer as coisas não é a melhor forma de trabalhar em colaboração, mas provavelmente continua a ser feita assim por causa do conformismo, e de alguma preguiça que puxa sempre para fazer as coisas desta forma porque é mais confortável. E depois este tipo de projetos tem uma visão burocrática de produção de informação, que está muita orientada para os deliverables [...] e para uma avaliação tipo checklist, muito burocrata. Como as propostas funcionam como uma espécie de contrato, entre financiador e parceiros, este contexto faz com que a melhor forma de organizar a informação seja em pastas de acordo com Work Packages.

E essa organização também permite que os colaboradores saibam exatamente onde estão os conteúdos que precisam.

Anónimo 3: Pois, e não estou a ver muitas alternativas a este tipo de organização de conteúdos. Se estivermos a falar da organização dos conteúdos no decurso do projeto, acho que essa forma de organizar é a que se adequa melhor à gestão de processos do projeto, que estão organizadas em work packages. Agora, se estivermos a pensar em utilizar essa informação pós-projeto, penso que já não é adequado.

Pois, porque perde-se a relação da informação com os processos do contexto.

Anónimo 3: Sim. Apesar do Plone ter live search e indexação automática dos conteúdos, isso não é suficiente. Se quisermos explorar, passado um ano, os conteúdos dos projetos, penso que não faz muito sentido estar a pesquisar por “work packages”, se calhar faz mais sentido por temática ou assunto.

Pois. O problema é que mesmo as opções de categorização do Plone não são utilizadas pelos colaboradores.

Anónimo 3: Pois não. E essas questões já acabam por cair na área do Comportamento Informacional.

[...]

Anónimo 3: A vertente também que eu pretendo na gestão da informação dos projetos, é também passar um pouco da gestão documental para a gestão de conteúdos, para a colaboração e para a discussão. Isto hoje em dia, quando é preciso discutir um tópico qualquer, é feito através de muitos e-mails, e eu tentei que isso fosse feito de forma no sistema de gestão de conteúdos. Por exemplo, num fórum.

Isso é estranho porque, apesar da submissão e armazenamento de conteúdos ser menos amigável no Plone que, por exemplo, na Dropbox, o acesso ao fórum requer o mesmo esforço que a consulta do e-mail.

Anónimo 3: Pois. O problema é que as pessoas estão habituadas a centrar a sua atividade à volta do e-mail, e não querem ser forçadas a utilizar várias plataformas. No entanto, é importante que se comece a perceber que uma plataforma universal nunca existirá.

[...]

Anónimo 3: É lógico que um utilizador que chega ao sistema de gestão de conteúdos de um projeto não irá perceber o domínio do conhecimento através das work packages. Se ele pudesse aprender com a taxonomia desse site, ou mesmo com a informação registada no fórum, seria muito melhor.

[...]

Anónimo 3: É importante que haja uma divisão do esforço, entre tarefas de indexação que podem ser automatizadas e tarefas de classificação por parte dos colaboradores, e que haja um investimento na criação de ajudas à classificação que facilitem o trabalho o mais possível às pessoas, que não as perturbem mesmo. É importante que se motive uma cultura de utilização da classificação. Isto também está relacionado com a própria funcionalidade e usabilidade dos sistemas, claro.