



Perspectivas de investigación

Profissionais da informação e medievalistas em torno da construção de uma taxonomia de História Medieval Portuguesa: dificuldade e reptos

Filipa Medeiros

Centro de História, Culturas e Sociedades,
Universidade de Évora
Portugal · fmcmm@uevora.pt

José Antonio Moreiro González

Universidad Carlos III de Madrid,
Departamento de Biblioteconomía y Documentación
España · jamore@bib.uc3m.es

Resumo: O objetivo principal desta comunicação consiste em apresentar uma taxonomia de História Medieval Portuguesa (sécs. XII-XV), elaborada no âmbito do Doutoramento em Ciências da Informação e Documentação da Universidade de Évora, em colaboração com o Departamento de Biblioteconomía y Documentación da Universidad Carlos III de Madrid. Trata-se de uma investigação aplicada, caracterizada por uma abordagem interdisciplinar do objeto de estudo, e que procura responder à necessidade de indexação de recursos digitais, em particular os contidos em bases de dados especializadas, consideradas hoje veículos fundamentais para a disseminação e partilha da produção científica em acesso aberto. Para a construção da taxonomia utilizaram-se os métodos já comprovados com êxito noutros estudos de caso para desenhar, desenvolver e manter vocabulários controlados deste tipo. Por conseguinte, utilizou-se o método descendente ou do geral para o particular (*top-down*). Como principais conclusões, importa referir que a taxonomia criada, tratando-se de uma investigação aplicada, testou com êxito os métodos anteriormente comprovados por especialistas para desenvolver vocabulários controlados desta natureza. Sublinha-se, ainda, o caráter pioneiro da taxonomia no panorama dos estudos medievais portugueses. A sua elaboração, exatamente pela ausência de estruturas comparativas ao nível nacional e internacional, constituiu um enorme repto, só alcançado pelo trabalho interdisciplinar entre medievalistas e profissionais da informação. Por isso mesmo, reforça-se a ideia de que a taxonomia elaborada representa, acima de tudo, uma primeira tentativa de categorização terminológica deste domínio disciplinar. Enquanto tal, ela é discutível, encontrando-se em constante evolução formal e semântica, estando os seus termos a ser testados em bases de dados especializadas. Por fim, resta mencionar que esta taxonomia tem essencialmente uma natureza terminológica, constituindo, por isso, o seu desenvolvimento ontológico um dos seus principais desafios futuros.

Palavras-chave: Representação da informação; Vocabulários controlados para a *web*; Taxonomias; História Medieval Portuguesa.

Abstract: The main goal of this communication is to present a taxonomy of Portuguese Medieval History (13th-15th centuries), prepared within the scope of the Doctoral Programme in Library and Information Science of the Universidade de Évora, in collaboration with the Departamento de Biblioteconomía y Documentación of the Universidad Carlos III de Madrid. This is an applied research work characterized by an interdisciplinary approach to the object of study that seeks to meet the needs for the indexation of digital resources, namely the ones contained in specialized databases, which are currently seen as crucial vehicles for disseminating and sharing open-access scientific knowledge. In order to build this taxonomy we used methods that have already proven successful in other case studies to design, develop and maintain this type of controlled vocabularies. So, we used a top-down, or general-to-specific, method. As main conclusions, we should mention that the taxonomy that was created, being the result of an applied research work, has successfully tested the methods that had been previously proven by experts for the construction of this type of controlled vocabularies. We

should also highlight the pioneering nature of this taxonomy within the field of Portuguese Medieval Studies. Precisely due to the fact that there were no national or international comparative frameworks, its preparation was a great challenge that we were only able to meet thanks to an interdisciplinary collaboration between medievalists and I&D professionals. That is why we would like to stress the idea that the taxonomy we have built is, above all, a first attempt at defining a terminological categorization within this subject matter. As such, it is open for discussion and undergoes a constant formal and semantic evolution process, and its terms are being tested in specialized databases. Finally, we should mention that this taxonomy has an essentially terminological nature and, therefore, its ontological development shall be one of its main future challenges.

Keywords: Representation of information; Controlled vocabularies for the web; Taxonomies; Portuguese Medieval History.

1. Introdução

Atualmente, a interdisciplinaridade assume-se como um elemento fundamental da investigação científica nos mais diversos domínios disciplinares. Nas Ciências da Informação e Documentação (CID), designadamente no âmbito da representação da informação e da construção de vocabulários controlados especializados, a visão interdisciplinar torna-se ainda mais indispensável dado o elevado grau de especificidade da terminologia utilizada em cada uma das áreas científicas.

Por conseguinte, este estudo coloca em interação duas áreas do saber tradicionalmente alheias uma da outra: as CID e a História Medieval.

A elaboração de uma taxonomia no âmbito da História Medieval constitui, como adiante se verá, uma opção pioneira e como tal sem parâmetros passíveis de comparação.

Não obstante as dificuldades e as opções, sempre discutíveis, subjacentes à sua elaboração, a taxonomia constitui um instrumento de inegável interesse e de apoio aos investigadores, em especial numa sociedade marcada pelo crescimento exponencial da informação disponível em acesso aberto.

Ela deve ser entendida como um instrumento em permanente construção e avaliação, mas clara nos propósitos que se propõe atingir.

2. Objetivos

O objetivo deste estudo é aplicar os métodos comprovados anteriormente por autores da especialidade para construir uma taxonomia destinada a organizar e a representar informação sobre História Medieval Portuguesa, relativa aos séculos XII-XV, já disponibilizada em ambiente *web* sob a forma de listagem, como se apresentará mais à frente.

Sobre a taxonomia, efetua-se, *a priori*, duas importantes ressalvas: uma, de ordem cronológica; outra, de ordem temática. A primeira ressalva relaciona-se com o facto de a taxonomia englobar o período que medeia entre o início do reinado de D. Afonso Henriques (1143) e o final do reinado de D. João II (1495), baliza temporal que corresponde ao estabelecido nas obras de referência da especialidade (Serrão e Marques, 1987 e 1996) e que toma como base a criação do reino de Portugal e a sua evolução política. A segunda ressalva diz respeito ao contexto temático abrangido pela taxonomia. Deste modo, considerou-se a História Medieval nas suas diferentes vertentes de análise, mas colocaram-se de parte categorias temáticas que, por um lado, incidem sobre cronologias mais avançadas – de que é exemplo a História dos Descobrimentos – ou, por outro, se ocuparem de objetos de estudo muito especializados, tais como a Filologia, Arqueologia e Epigrafia, História Literária, História da Arte e História da Música. Tal não significa que, posteriormente e no decurso do desenvolvimento da taxonomia, não sejam incluídas essas categorias temáticas.

Note-se, ainda, que não se aprofundaram as realidades muçulmana e judaica, tendo em conta a especificidade destes domínios temáticos (ex. áreas da cultura,

religião, entre outras). Não obstante, o seu desenvolvimento deverá constituir-se como um dos objetivos futuros no contexto da ampliação da taxonomia.

Concetualmente, uma taxonomia é entendida como um vocabulário controlado, organizado de forma hierárquica, no qual se estabelecem relações semânticas simples (generalização/especialização) entre os termos que o compõem (Conway e Sligar, 2002; Gilchrist, 2003; Zhonghong e outros, 2006; Moreiro, 2011). Devido ao seu elevado grau de especialização, a taxonomia em observação trata-se de uma «taxonomia de domínio» ou «taxonomia de especialidade» (Campos e Gomes, 2007: 4).

A criação desta taxonomia procura colmatar uma lacuna existente no plano das unidades de informação portuguesas especializadas em História Medieval, muitas delas incorporadas nas universidades e respetivos centros de investigação. Na verdade, é premente o tratamento diferenciado da informação da especialidade, designadamente no que diz respeito à indexação do universo temático da História medieval portuguesa. Para além disso, a elaboração deste vocabulário controlado procura, sobretudo, responder à necessidade de indexação de recursos digitais, nomeadamente os contidos em bases de dados especializadas, considerados hoje veículos fundamentais para a disseminação e partilha da produção científica em acesso aberto. Veja-se, a título exemplificativo, algumas bases de dados já produzidas pelo Instituto de Estudos Medievais da Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa (IEM - FCSH/UNL) e disponíveis em: <http://iem.fcsh.unl.pt/section.aspx?kind=noticia&id=49>

Resta mencionar que esta investigação resulta de parte de uma tese de doutoramento desenvolvida na Universidade de Évora, Portugal, (2009-2014), em colaboração com o Departamento de Biblioteconomía y Documentación da Universidade Carlos III de Madrid (Medeiros, 2014).

3. Metodologia

A origem da inferência no interior de uma taxonomia reside na sua dimensão hierárquica em níveis de dependência terminológica. Deste modo, é possível subir até ao nível mais genérico ou descer até ao nível mais específico, passando pelos diferentes níveis intermédios da representação da realidade. Por conseguinte, uma taxonomia categoriza os termos autorizados pela experiência e seus possíveis atributos dentro de um domínio concreto para representar os seus conteúdos e as suas respetivas relações hierárquicas, estabelecendo-se, portanto, como especificação explícita de uma representação cognitiva que inclui uma lista estruturada de termos (Cumming, 2003: 3). No fundo, as taxonomias explicitam um modelo conceitual de domínio, permitindo, de igual modo, agregar dados.

As taxonomias sempre funcionaram como estruturas conceituais e de classificação para organizar e recuperar entidades de informação. Na atualidade, aplicam-se para ordenar recursos digitais alojados em servidores Web, de molde a categorizá-los e a navegar sobre eles. Converteram-se, assim, em meios de acesso, mais com uma função de visualização do que recuperação, pois ajudam os utilizadores a encontrar produtos e serviços na Web, assim como a representar melhor as suas necessidades de informação (Moreiro, 2006: 97). Atuam, pois, como mapas conceituais dos assuntos sobre um campo especializado.

Uma das suas mais-valias reside na possibilidade de navegação por categorias, que conduz a uma progressiva motivação por parte dos utilizadores, que já não necessitam de possuir um conhecimento profundo desse domínio, pelo contrário, ele vai sendo descoberto e explorado por meio da navegação. Por conseguinte, as taxonomias constituem-se como verdadeiros guias de assunto a ser investigado através de processos de descoberta de informação (Campos e Gomes, 2007: 3-4). Contudo, as taxonomias também apresentam algumas debilidades que devem ser tomadas em consideração. Por exemplo, na maioria dos casos, as taxonomias contêm somente relações hierárquicas. Consequentemente, no caso de se pretender incluir outro tipo de relações, deverão ser complementadas com mecanismos de pesquisa adicionais (Moreiro, 2011: 62).

Tendo em conta os aspetos referidos anteriormente e, sobretudo, a necessidade de indexar recursos digitais de informação ligados ao universo dos estudos medievais portugueses, optou-se pela criação de uma taxonomia por todas as potencialidades que este tipo de vocabulário controlado apresenta em ambiente *web*.

Para a construção e desenvolvimento da taxonomia utilizou-se o método descendente ou do geral para o particular (*top-down*) (Jagerman, 2006), o mais utilizado na criação deste tipo de vocabulário controlado. Este método compõe-se das seguintes etapas:

- a) Compilar o conhecimento;
- b) Reduzir os sinónimos e proceder à escolha dos termos preferenciais;
- c) Preparar uma estrutura sistemática prévia;
- d) Elaborar uma primeira versão do esquema;
- e) Completar a taxonomia com todos os seus termos;
- f) Avaliar o seu funcionamento;
- g) Publicar uma primeira versão.

Na fase prévia à construção da taxonomia, observaram-se dois princípios básicos (Moreiro, 2011: 55):

- Dimensão: como já se referiu, a elaboração da taxonomia procurou responder ao problema da indexação de informação especializada sobre História medieval portuguesa, em particular a existente em suporte digital. Como tal, pretende fornecer uma lista de termos controlados que permitam ora auxiliar o indexador na tarefa de representação dos assuntos, ora o utilizador na pesquisa e recuperação da informação útil e pertinente. Estima-se que a taxonomia incorpore progressivamente novos termos, à medida da sua crescente utilização e desenvolvimento, estando, pois, preparada para o efeito. Este processo deverá ser acompanhado pela consulta de especialistas, que deverão trabalhar em conjunto com os profissionais I&D, em particular na tarefa de validação terminológica. Por último, prevê-se que a manutenção desta estrutura se concretize em suporte informático (*software* de taxonomias), disponível em ambiente *web*, aspeto que se prevê ser uma das suas mais-valias futuras, assim como a agregação de recursos nascidos digitais.
- Termos a utilizar: na fase de recolha dos termos, foi dada especial atenção à questão da comunicabilidade, por outras palavras, que os conceitos em linguagem natural utilizados pela comunidade de especialistas correspondessem, sempre que possível, aos termos em linguagem controlada constantes da taxonomia. Esta opção apresenta como objetivo motivar os utilizadores a navegar pela estrutura e a pesquisar os seus termos, tarefa facilitada pelo trabalho de compatibilização entre a linguagem natural e o vocabulário controlado. Além disso, atendeu-se à utilidade dos termos, o que levou a selecionar só os conceitos com representatividade semântica no contexto da historiografia medieval portuguesa. A perceção do peso de cada assunto no âmbito da produção científica sobre História medieval portuguesa foi facilitada pela importância que as fontes de informação específicas dão a cada um dos assuntos.

Não obstante, estas tarefas revelaram-se sempre um enorme desafio, implicando escolhas, logicamente discutíveis à luz da comunidade de medievalistas. Por isso mesmo, a taxonomia que se apresenta – inédita no panorama do medievalismo português – deve ser entendida, antes de mais, como uma primeira proposta de sistematização terminológica desta área do saber, encontrando-se em aberto, não sendo, portanto, um fim em si mesma.

Apresenta-se, de seguida, o desenvolvimento específico de cada uma das etapas da construção da taxonomia de História Medieval Portuguesa, de acordo com o método descendente:

a) **Compilação do conhecimento:** nesta primeira etapa procedeu-se ao levantamento e compilação dos assuntos a incorporar na estrutura classificatória hierárquica e que constituirão os pontos de acesso à informação. Este levantamento foi efetuado a partir de várias formas, entre as quais: instrumentos classificatórios ou terminológicos já existentes; fontes de informação da especialidade (vocabulários, léxicos, dicionários, entre outros); e consultas a especialistas. Assim, e num primeiro momento, realizou-se o mapeamento dos departamentos, institutos/centros de investigação e projetos nacionais e estrangeiros sobre Idade Média, com vista ao seu contacto (via correio eletrónico) para averiguar se dispunham de vocabulários controlados que secundassem a indexação dos seus acervos bibliográficos e da sua produção científica impressa e/ou digital. Para tal, recorreu-se, em grande parte, à seleção de *links* sugeridos no sítio *web* do IEM - FCSH/UNL, 25 na sua totalidade, seleção esta disponível em: <http://iem.fcs.unl.pt/section.aspx?kind=noticia&id=51>. Dos 25 pedidos de colaboração efetuados, obtiveram-se 11 respostas. Quase todas elas apontaram a inexistência de vocabulários controlados da especialidade criados especificamente para indexar informação relativa à Idade Média. Porém, as respostas recolhidas demonstraram um manifesto interesse na investigação, solicitando alguns dos respondentes, inclusive, a posterior disponibilização da taxonomia, entendendo-a como uma ferramenta essencial à melhoria do tratamento técnico no domínio dos estudos medievais em geral e, também, à facilitação da pesquisa e recuperação da informação, incrementada pela possibilidade de navegação e de seleção dos termos por parte dos próprios utilizadores. A exceção partiu do *Groupe d'Anthropologie Historique de l'Occident Médiéval (GAHOM)*, que nos comunicou o desenvolvimento de alguns tesouros temáticos, em particular o *Thesaurus Exemplorum Medii Aevi (ThEMA)* e o *Thesaurus des Images Médiévales en Ligne (TIMEL)*.

Num segundo momento, procedeu-se ao levantamento das fontes de informação da especialidade, atendendo sobretudo aos vocabulários, léxicos e dicionários sobre Idade Média. Não foram encontradas quaisquer produções de origem portuguesa destas tipologias documentais, pelo que se reuniram contribuições estrangeiras de referência no âmbito da História Medieval geral e que poderiam enriquecer a taxonomia, particularmente no que se relaciona com a constituição das suas categorias gerais, quer do ponto de vista da forma, quer do conteúdo. Chegou-se a um total de 11 fontes de informação estrangeiras. No que respeita às fontes de informação portuguesas, na impossibilidade de se dispor desse tipo de fontes em língua nacional, recorreu-se a histórias gerais e a manuais universitários de referência no âmbito dos estudos medievais portugueses, em concreto às seguintes fontes de informação: Serrão e Marques, 1987 e 1996; Marques, 1988; Tavares, 1990; Serrão, 1994; Coelho, 1991; Tavares, 1992; Moreno, 1995; Mattoso, 1997; Branco e Costa, 1992; Ramos e outros, 2009.

Por fim, e num terceiro momento, importa mencionar que o processo de construção da taxonomia foi constantemente acompanhado e monitorizado por um medievalista português e por um especialista em representação da informação, de resto os orientadores da investigação doutoral que, em parte, se apresenta nesta comunicação. Esta supervisão especializada concretizou-se na elaboração de várias propostas classificatórias, até se alcançar a versão final. Sempre que se revelou necessário, nomeadamente em domínios muito particulares da História Medieval, recorreremos a pareceres de outros especialistas que contribuíram para a clarificação e enriquecimento semântico da taxonomia.

b) **Redução dos sinónimos e escolha dos termos preferenciais:** este ponto corresponde à segunda etapa de construção da taxonomia, no qual se expõem as questões atinentes ao controlo formal e semântico do vocabulário, de acordo com o seguinte esquema de análise:

- Normalização formal: controlo morfológico (língua, género e número) e controlo sintático (termos simples e termos compostos);
- Normalização semântica: controlo das ambiguidades via significante (homógrafos ou polissemos) e via significado (palavras sinónimas e quase-sinónimas), através do estabelecimento de relações semânticas.

Normalização formal: uma vez que a língua de partida da taxonomia é a portuguesa, seguiu-se, naturalmente, as orientações dos normativos nacionais, em concreto a Norma Portuguesa 4036 (pontos 6 e 7) e o manual *Siporbase: sistema de indexação em português* (secção 4, Terminologia) (NP 4036, 1992: 8-22; Portugal, 1998). Veja-se então:

Controlo morfológico

- **Língua e ortografia:** A língua selecionada para a elaboração da taxonomia foi a língua portuguesa, em conformidade com o acordo ortográfico em vigor. Tal não invalida que, posteriormente, os termos não sejam traduzidos para outros idiomas, o que dependerá da aceitação e do impacto que esta ferramenta tenha nas unidades de informação da especialidade, ao nível nacional e internacional. Adotou-se a ortografia do uso corrente, preferindo os termos mais generalizados entre a comunidade científica e os utilizados nas fontes de informação de referência no domínio dos estudos medievais. Na construção da taxonomia, convencionaram-se ainda alguns aspetos relacionados com a ortografia, tais como: utilização de maiúscula na inicial dos termos simples e na inicial do primeiro elemento dos termos compostos; os acentos foram sempre utilizados, respeitando as normas da ortografia atual; recorreu-se ao uso de símbolos ortográficos, em particular ao parêntesis curvo (...) para construir os qualificadores dos termos homógrafos e, também, ao hífen (-) quando este é parte integrante de nomes comuns.
- **Género:** Os termos selecionados encontram-se quer no género masculino, quer no género feminino, de acordo com as normas estabelecidas na língua portuguesa. Nos casos em que um termo apresenta uma significação genérica, recorreu-se ao género masculino.
- **Número (escolha da forma singular ou plural):** no caso da taxonomia em estudo, adotámos, regra geral, a forma plural, na medida em que a maioria dos termos selecionados se integra na categoria de entidades concretas enumeráveis. Recorreu-se também à forma singular, designadamente nos casos de representação de entidades abstratas.

Controlo sintático

- **Termos simples e termos compostos:** Os termos compostos constantes da taxonomia foram mantidos na forma da linguagem natural. Por outras palavras, considerou-se a forma do uso corrente utilizada pelos medievalistas, com o intuito de facilitar a tarefa de pesquisa e de recuperação da informação. Quanto à ordem das palavras nos termos compostos, seguimos as orientações da Norma, inserindo-os, portanto, na sua ordem natural e não de forma invertida. Por fim, a tipologia dos termos relativamente ao seu conteúdo. Na taxonomia em observação, fez-se uso de termos temáticos, onomásticos, cronológicos e geográficos:
Ex. Devoções religiosas [termo temático]
Ex. Mafalda de Mouriana [termo onomástico]
Ex. Séc. 13 [termo cronológico]
Ex. Braga [termo geográfico]

Normalização semântica: o controlo semântico irá garantir, por um lado, a normalização conceitual e, por outro, a superação da ambiguidade característica da linguagem natural, geradora ora de silêncio, ora de ruído na recuperação da informação. Para tal, é necessário que a cada palavra corresponda um só significado. O controlo semântico consiste num procedimento que subentende dois fundamentos: um de natureza teórica e outro de natureza prática (Simões, 2008: 103-105). Na taxonomia em análise, no que se refere aos fundamentos de natureza teórica, tivemos em conta a principal finalidade da construção deste vocabulário controlado que é a indexação, pesquisa e recuperação de informação sobre História medieval portuguesa entre os séculos XII e XV, em ambiente *web*. Nesta perspetiva, foi reunido um conjunto de termos expressivos deste domínio temático, com recurso às fontes da especialidade, tendo sobretudo em consideração as necessidades de informação do

público-alvo da taxonomia, isto é, a comunidade científica e académica de medievalistas portugueses.

Quanto aos fundamentos de natureza prática, em particular ao nível do controlo semântico via significante, durante a fase de seleção dos termos, foi dado especial cuidado aos homógrafos ou polísemos, isto é, às palavras constituídas pela mesma cadeia de caracteres mas que têm significados diferentes. No entanto, estas palavras não se revelaram uma dificuldade efetiva, pois estamos perante um vocabulário controlado altamente especializado, no qual este tipo de situações é substancialmente menor se compararmos com outros vocabulários de cobertura temática de maior abrangência. Nestes casos, recorreremos ao uso dos designados «qualificadores» para contextualizar e fixar o âmbito semântico (significado) do termo e que, de resto, são parte integrante do termo de indexação, formando, assim, um termo composto:

Ex. Jantar (tributo senhorial)
Jantar (refeição)

Ao nível do controlo semântico via significado, ao reunirmos uma primeira amostra de termos da especialidade, efetuámos, à partida, a eliminação de sinónimos e de quase-sinónimos, elegendos termos preferenciais, destinados a representar univocamente um só conceito:

Ex. Cantigas de escárnio e de maldizer [termo preferencial]
Ex. Canções de escárnio e de maldizer [termo não preferencial]

O controlo semântico efetiva-se também através das relações semânticas existentes num vocabulário controlado. Desta forma, as taxonomias abrangem, na grande maioria dos casos, somente relações hierárquicas que constituem o seu eixo estruturante (ISO 25964-2, 2011: 68-69). Por conseguinte, a possibilidade de visualização e de navegação por toda a estrutura conceitual hierárquica da especialidade proporciona ao utilizador optar por duas formas de pesquisa – uma genérica e outra específica – que seleciona mediante as necessidades de informação do momento. Esta possibilidade é, de igual modo, de grande utilidade para o indexador, que poderá aceder facilmente a um termo mais genérico ou mais específico no ato de representação do conteúdo de um documento, não perdendo de vista, naturalmente, os interesses dos utilizadores e a política de indexação do serviço. Por último, as relações hierárquicas permitem o controlo semântico do vocabulário, conseguido pelo estabelecimento de termos genéricos e de termos específicos (Simões, 2008: 118).

No caso da taxonomia em observação, que apenas abrange relações hierárquicas, estas estabeleceram-se de acordo com os seguintes critérios:

- Regra geral, cada um dos termos apenas se encontra subordinado a um termo genérico. Todavia, e excecionalmente, existem alguns termos que pertencem a várias categorias, o que concorre para o facto de a taxonomia possuir relações não só mono-hierárquicas mas também relações poli-hierárquicas, mesmo que pontuais. Isto sucede, portanto, quando um termo, por razões lógicas, se encontra subordinado, ao mesmo tempo, a mais do que um termo genérico, como é o caso de alguns cargos administrativos:
Ex. Procuradores [Oficiais da administração central]
Ex. Procuradores [Oficiais concelhios]
- Existem termos que não estão subordinados a nenhum outro, constituindo-se como termos independentes (termos de topo);
- Os termos podem ter dois ou mais termos de hierarquia inferior à sua, sendo esta situação, aliás, a mais frequente nas categorias estabelecidas;
- A ordenação dos termos na estrutura hierárquica foi efetuada por ordem alfabética, exceto na categoria «Acontecimentos», «Personalidades» e «Reinados», nas quais os termos foram inseridos por ordem cronológica. Só deste modo, se conseguiu organizá-los de forma diacrónica, possibilitando ao utilizador navegar na estrutura taxonómica de acordo com a sucessão natural dos acontecimentos, das personalidades e dos reinados, respetivamente:

Ex. 1128, Batalha de São Mamede
Ex. 1350-1405, Leonor Teles
Ex. 1385-1433, Reinado de D. João I

c) Processo de construção da taxonomia: Após a normalização formal e semântica do vocabulário, conforme os critérios e procedimentos supramencionados, efetuou-se o registo alfabético dos termos compilados, distribuindo-os pelas categorias que compuseram a estrutura sistemática prévia da taxonomia, correspondendo esta operação à terceira etapa.

Seguidamente, e numa quarta etapa, elaborou-se uma primeira versão do esquema taxonómico, definindo a posição hierárquica dos termos dentro das suas respetivas categorias e ordenando, por conseguinte, cada termo específico sob o termo genérico correspondente. Por sua vez, no interior de cada categoria, os conceitos foram organizados em classes, que podem ser de dois tipos: cadeias e renques. As primeiras são séries verticais de conceitos e as segundas séries horizontais de conceitos, podendo ser ambas do tipo genérico (é tipo de?) ou partitivo (é parte de? é todo de?) (Campos e Gomes, 2007: 6).

Estes procedimentos permitiram avançar para a quinta etapa e elaborar a superestrutura classificatória da taxonomia e completá-la com todos os seus termos. A sua versão final é constituída por 17 categorias, a saber:

- 1) Administração
- 2) Áreas da História
- 3) Ciências e técnicas historiográficas
- 4) Coroa
- 5) Cronologia
- 6) Cultura
- 7) Economia
- 8) Estruturação social
- 9) Fontes
- 10) Geografia
- 11) Guerra
- 12) Personalidades
- 13) Recursos de informação
- 14) Reinados
- 15) Religião
- 16) Sociedade
- 17) Vida privada

O processo de construção da taxonomia revelou-se bastante moroso e exigiu grande minúcia no que respeita à seleção e à normalização formal e semântica dos termos a incorporar, assim como a sua respetiva ordenação na estrutura classificatória hierárquica. Esta tarefa viu a sua dificuldade acrescida devido à ausência de sistemas de organização do conhecimento do mesmo âmbito temático que pudessem servir como ponto de comparação.

Como escolha metodológica de base, decidiu-se elaborar a taxonomia de forma gradativa, isto é, categoria a categoria, arquitetando, a par e passo, as classes correspondentes e, dentro destas, as suas cadeias e renques.

No decurso do registo dos termos de cada uma das categorias, foram elaboradas listas complementares com outros termos candidatos a eventual incorporação no esquema final. Estas listas continham não apenas os casos de sinonímia, mas também termos sobre os quais havia dúvidas relativamente à sua representação conceitual ou à sua pertinência informativa.

Numa fase posterior, e depois de superados os casos de ambiguidade e de justificação da existência de determinados termos, procedeu-se ao estabelecimento de todas as categorias da versão final e completou-se a taxonomia com todos os seus termos preferentes, acautelando que nenhum conceito significativo ficasse omisso de representação.

Os termos integrantes do esquema taxonómico final foram todos cuidadosamente revistos, do ponto de vista formal e semântico, de modo a evitar futuras ambiguidades no ato de indexação e de pesquisa e recuperação da informação da especialidade.

Numa sexta etapa, avaliou-se a taxonomia de acordo com as recomendações da ANSI/NISO Z39.19-2005, pois quer a NP 4036, quer a ISO 25964 são omissas relativamente a esta questão. Deste modo, optou-se pela avaliação heurística da versão final da taxonomia, pois pensa-se ser aquela que mais se adequa ao vocabulário controlado construído, que é uma taxonomia de domínio, e que, enquanto tal, pressupõe a intervenção e monitorização efetivas e permanentes por parte de especialistas no processo de elaboração da mesma, como, de resto, ocorreu (ANSI/NISO Z39.19, 2005: 95-96).

Desta forma, foi solicitada a colaboração formal a dois especialistas, um da área da História medieval portuguesa e outro da área das CID, especializado em vocabulários controlados e cuja avaliação se apresenta mais adiante.

Na sétima e última etapa, publicou-se uma primeira versão da taxonomia. Os termos recolhidos e normalizados foram, na sua totalidade, introduzidos num *software* próprio de gestão de taxonomias denominado *Knowledge Manager* (KM), comercializado pela empresa espanhola *The Reuse Company*, que colabora com o Departamento de Biblioteconomia e Documentação da Universidade Carlos III de Madrid. A versão utilizada é uma versão para *download* de teste (edição 5.0.0.). O KM permite a criação de diferentes tipos de relatórios que podem ser exportados em vários formatos, possibilitando a introdução dos termos para outras aplicações.

4. Exposição dos resultados e sua discussão

A taxonomia conta atualmente com 2987 termos, em permanente revisão e atualização.

A versão integral da taxonomia encontra-se disponível online em: <http://www.cidehus.uevora.pt/Recursos/Taxonomia-de-Historia-Medieval-Portuguesa>

Tal como se depreende pelo gráfico 1, a taxonomia de História Medieval Portuguesa é constituída por 17 categorias, com diferentes níveis de profundidade:

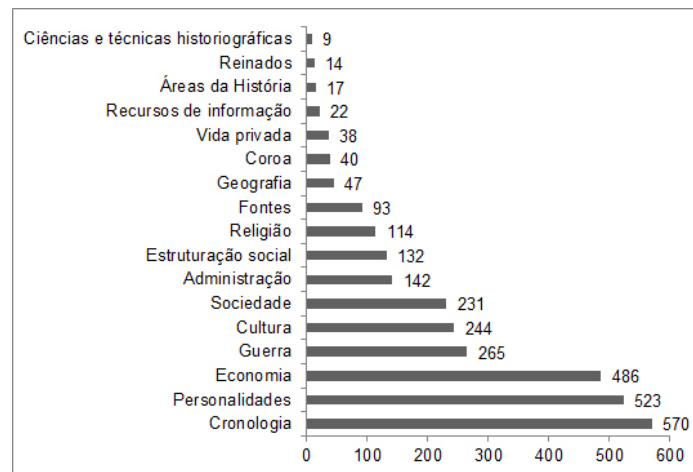


Gráfico 1: Distribuição do nº de termos pelas categorias da taxonomia de História Medieval Portuguesa

Assim, a categoria de maior extensão é a «Cronologia», que conta com 570 termos. Ao invés, a categoria «Ciências e técnicas historiográficas» reúne somente 9 termos.

De igual modo em posição de destaque, encontram-se as categorias «Personalidades», «Economia», «Guerra», «Cultura» e «Sociedade» com 523, 486, 265, 244 e 231 termos, respetivamente.

Numa hierarquia intermédia, listam-se as categorias «Administração», «Estruturação social», «Religião» e «Fontes», reunindo, na devida ordem, 142, 132, 114 e 93 termos.

Por fim, e com um menor número de termos, identificam-se as categorias «Geografia», «Coroa», «Vida privada», «Recursos de informação», «Áreas da História»,

«Reinados» e «Ciências e técnicas historiográficas», que agregam 47, 40, 38, 22, 17, 14 e 9 termos, respetivamente.

É notória a existência de diferentes níveis de profundidade das categorias do vocabulário taxonómico em análise. Esta situação deve-se ao facto das categorias deterem diferentes escopos de análise. Outro dos fatores relaciona-se com a distribuição desigual das fontes disponíveis para cada uma das categorias e a forma como os diferentes temas se encontram representados na produção científica universitária sobre História Medieval.

No que respeita à avaliação da taxonomia, em termos gerais, solicitou-se ao medievalista a avaliação semântica da taxonomia (atualidade e comunicabilidade dos termos) e ao especialista da área da representação da informação a avaliação formal (morfológica e sintática) da mesma. Para tal, foram elaboradas duas grelhas de análise, de acordo com as recomendações da ANSI/NISO Z39.19-2005. Os resultados de ambas as avaliações foram bastante positivos e funcionaram de forma complementar, tal como é inerente a qualquer investigação de índole interdisciplinar e que, no contexto em observação, se assume como fundamental pela necessidade de validação terminológica conjunta.

As tabelas 1 e 2 apresentam os critérios de avaliação delineados e a respetiva pontuação atribuída pelos especialistas a cada um deles:

<i>CRITÉRIOS DE AVALIAÇÃO</i>		<i>PONTUAÇÃO</i>
1.	Fontes de informação gerais e específicas utilizadas	Muito bom
2.	Uniformidade ao nível da representação das várias áreas temáticas incluídas na taxonomia	Muito bom
3.	Pertinência semântica dos termos face ao vocabulário utilizado pelos utilizadores (atualidade dos termos)	Muito bom
4.	Adequação da estrutura hierárquica	Muito bom
5.	Eficiência do sistema de navegação	Muito bom
6.	Eficácia do sistema de pesquisa (recuperação e opções de pesquisa)	Muito bom
7.	Observações/Sugestões:	
Solicitações de pareceres específicos a especialistas oriundos dos vários domínios temáticos abrangidos pela taxonomia.		

Tabela 1: Grelha de avaliação (medievalista)

<i>CRITÉRIOS</i>		<i>PONTUAÇÃO</i>
1.	Fontes de referência e procedimentos utilizados na normalização do vocabulário	Muito bom
2.	Consistência morfológica e sintática dos termos	Muito bom
3.	Correção concetual ao nível do estabelecimento das relações hierárquicas	Muito bom
4.	Eficiência do sistema de navegação	Muito bom
5.	Eficácia do sistema de pesquisa (recuperação e opções de pesquisa)	Bom
6.	Observações/Sugestões:	
Diversificação e refinamento do leque de pesquisas.		

Tabela 2: Grelha de avaliação (especialista em vocabulários controlados)

Para a versão final da taxonomia, preferiu-se a apresentação alfabética, pois é aquela que possibilita a navegação pela estrutura semântica do vocabulário taxonómico, permitindo a visão geral das suas categorias, assim como das suas respetivas cadeias e renques, tal como demonstra a figura 1:

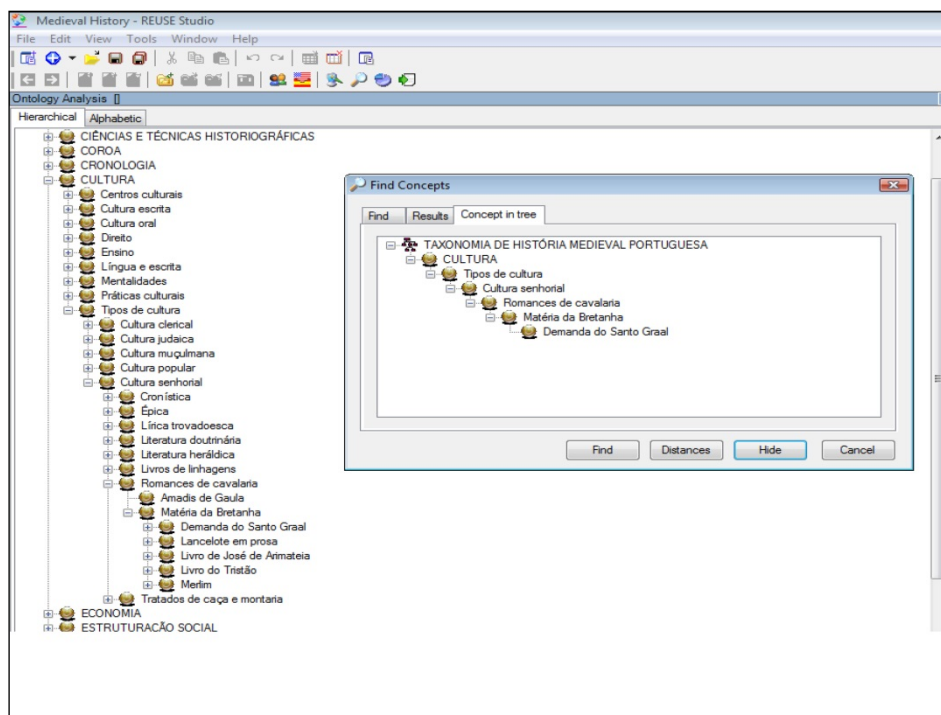


Figura 1: Visão geral da taxonomia de História Medieval Portuguesa no software KM

5. Conclusões

Em jeito de conclusão tecem-se algumas considerações fundamentais:

Em primeiro lugar, importa salientar que a taxonomia construída, enquanto uma investigação aplicada, testou com sucesso os métodos já experimentados por especialistas para desenhar, desenvolver e manter vocabulários controlados desta natureza.

Em segundo lugar, há que ter em linha de conta que embora este vocabulário taxonómico seja representativo de um país e de um momento histórico em particular, mesmo assim poderá funcionar como ponto de partida para a construção de outros vocabulários controlados com o mesmo escopo temático, mas relativos a outras geografias, com o mesmo recorte cronológico.

Sublinha-se, por fim, o caráter pioneiro deste vocabulário taxonómico no panorama dos estudos medievais portugueses. A sua elaboração, exatamente pela ausência de estruturas comparativas ao nível nacional e internacional, constituiu um enorme repto, só alcançado pelo trabalho interdisciplinar entre medievalistas e profissionais da informação. Por isso mesmo, reforça-se a ideia de que a taxonomia elaborada representa, acima de tudo, uma primeira tentativa de categorização terminológica neste domínio disciplinar. Enquanto tal, ela é discutível, encontrando-se os seus termos em constante evolução formal e semântica e, de igual modo, a serem testados atualmente em bases de dados especializadas.

6. Trabalhos futuros

No que diz respeito a trabalhos futuros, seria interessante incorporar termos relativos a cronologias mais recuadas da Idade Média portuguesa, nomeadamente ao período da Antiguidade Tardia e à Época Islâmica, e também aos inícios do século XVI. O mesmo é válido relativamente à incorporação das temáticas que ficaram de parte nesta investigação, como por exemplo a História da Arte, a História Literária ou a História dos Descobrimentos e, ainda, o aprofundamento das realidades judaicas e muçulmanas. A concretizar-se, este alargamento iria implicar a constituição de uma equipa interdisciplinar, ainda mais alargada, que trabalhe em conjunto com os

profissionais da informação, particularmente na tarefa de validação terminológica, tendo em conta o elevado grau de especialização destas áreas do saber.

A agregação de recursos digitais, tais como fontes, bibliografias ou outros repertórios da especialidade, seria também uma mais-valia para o enriquecimento da taxonomia.

Para além disso, e numa tentativa de colocar à disposição de uma comunidade mais alargada de utilizadores, seria muito útil a tradução da taxonomia para outros idiomas, podendo assim servir de ponto de partida para a elaboração de outros vocabulários controlados ou, até mesmo, para o aperfeiçoamento deste.

Um outro trabalho futuro poderá ser a criação de um repositório temático de produção científica sobre História Medieval portuguesa, integrando o KM enquanto ferramenta de pesquisa e de recuperação da informação da especialidade. Respondendo ao atual desafio do acesso aberto e da disseminação e partilha do conhecimento, crê-se que um recurso deste género serviria não só para aproximar as diferentes universidades / centros de investigação portugueses com produção científica em História medieval, como também para aumentar a visibilidade deste campo disciplinar no contexto da ciência nacional, pelo menos no domínio das Ciências Sociais e Humanas, das quais se encontra mais próximo.

Por fim, deve procurar-se gerir o conhecimento contido na taxonomia com um maior alcance, isto é, para além daquele contido na organização hierárquico-classificatória do seu vocabulário. Pretende-se, assim, orientar o vocabulário taxonómico para uma esquematização que permita ligar os assuntos da investigação que representa a outros vocabulários existentes e com os quais detenha afinidades. Para alcançar esta interoperabilidade, terá que se estabelecer esquemas SKOS que codifiquem a taxonomia em XML e que permitam a sua migração no âmbito da web semântica (Daconta e outros, 2003). Desta forma, este vocabulário taxonómico é, ainda, fundamentalmente terminológico, não detendo, por isso, uma natureza ontológica. Por conseguinte, e numa fase mais avançada, este será, sem dúvida, o grande desafio futuro que se coloca ao vocabulário taxonómico aqui apresentado.

7. Agradecimentos

Agradece-se aos medievalistas e profissionais da informação a sua colaboração na validação semântica e formal dos termos da taxonomia, respetivamente.

Este artigo teve ainda a colaboração da empresa espanhola *The Reuse Company*, que cedeu gratuitamente o *software* KM, que possibilitou a elaboração da taxonomia e a publicação da sua primeira versão.

8. Referências bibliográficas

Nota: Na impossibilidade de aqui apresentar todas as referências bibliográficas que constituíram as nossas fontes de informação para a obtenção de dados, optámos por incluir apenas aquelas que dizem respeito a pontos de situação da historiografia medieval portuguesa e, também, algumas obras de referência da especialidade que serviram de base ao estabelecimento das temáticas, das cronologias e de determinados conceitos.

Referências bibliográficas sobre História Medieval

Branco, Maria João Violante; Costa, Adelaide Millán da (1992). *História da Idade Média: caderno de apoio*. Lisboa: Universidade Aberta. ISBN 972-674-370-2.

Coelho, Maria Helena da Cruz (1991). *História medieval de Portugal: guia de estudo*. Porto: Universidade Portucalense. ISBN 972-9354-07-3.

Marques, A. H. de Oliveira Marques (1988). *Guia do estudante de História medieval portuguesa*. 3.^a ed. Lisboa: Estampa.

Mattoso, José (dir.) (1997). *História de Portugal: a monarquia feudal (1096-1480)*. Lisboa: Estampa, vol. 2. ISBN 972-33-1263-8.

Morero, Humberto Baquero (1995). *História de Portugal medievo: político e institucional*. Colab. Maria da Conceição Falcão Ferreira, Luís Carlos do Amaral, Luís Miguel Duarte. Lisboa: Universidade Aberta, 2 vols. ISBN 972-674-132-7 (vol. 1). ISBN 972-674-135-1 (vol. 2).

Ramos, Rui; Sousa, Bernardo Vasconcelos e; Monteiro, Nuno Gonçalo (2009). *História de Portugal*. Lisboa: A Esfera dos Livros. ISBN 978-989-626-139-9.

Serrão, Joaquim Veríssimo (1994). *História de Portugal: Estado, pátria, nação*. 5.^a ed. Lisboa: Verbo, vol. 1. ISBN 972-22-0271-5.

Serrão, Joel; Marques, A. H. de Oliveira (dir.) (1987). *Nova História de Portugal: Portugal na crise dos séculos XIV e XV*. Coord. A. H. de Oliveira Marques. Lisboa: Estampa, vol. 4.

Serrão, Joel; Marques, A. H. de Oliveira (dir.) (1996). *Nova História de Portugal: Portugal em definição de fronteiras (1096-1325): do Condado Portucalense à crise do século XIV*. Coord. Maria Helena da Cruz Coelho, Armando Luís de Carvalho Homem. Lisboa: Presença, 1996. ISBN 972-23-2039-4, vol. 3.

Tavares, Maria José Ferro (1990). *Sociedades e culturas portuguesas*. Lisboa: Universidade Aberta, 2 vols. ISBN 972-674-030-4 (vol. 1). ISBN 972-674-049-5 (vol. 2).

Tavares, Maria José Ferro (1992). *História de Portugal medieval: economia e sociedade*. Lisboa: Universidade Aberta. ISBN 972-674-100-9.

Referências bibliográficas sobre representação da informação:

ANSI/NISO Z39.19-2005. *Guidelines for the construction, format and management of monolingual controlled vocabularies*. NISO Press: Bethesda, MD. ISBN 1-880-124-65-3.

Campos, Maria Luiza; Gomes, Hagar Espanha. Taxonomia e classificação: a categorização como princípio. *Enancib: Encontro Nacional de Pesquisa em Ciência da Informação*. [Electronic Version]. Salvador, 8, 2007. Retrieved 22 sep. 2010, from <http://www.enancib.ppgci.ufba.br/artigos/GT2--101.pdf>

Conway, Susan; Sligar, Char (2002). *Unlocking knowledge assets*. Redmont: Microsoft Press. ISBN 0-7356-1463-6.

Cumming, Maewin (2003). Tomatoes are not the only fruit: a rough guide to taxonomies, thesauri, ontologies and the like. *Record Management Bulletin*, n° 113, 1-4.

Daconta, M.; Obrst, L. J.; Smith, K. T. (2003). *The semantic web. A guide to the future of XML, web services and knowledge management*. Indianapolis: Wiley.

Gilchrist, Alan (2003). Thesauri, taxonomies and ontologies: an etymological note. *Journal of documentation*, vol. 59, (1) 7-18. ISSN 0022-0418.

Jagerman, Evert (2006). *Creating, maintaining and applying taxonomies*. Zoetermeer: E. Jagerman. ISBN 90-811-274-1-1.

ISO 25964-2011. Information and Documentation. *Thesauri and interoperability with other vocabularies. Part 1: Thesauri for information retrieval. Part 2: Interoperability with other vocabularies*. Geneva: International Organization for Standardization.

Medeiros, Filipa (2014). *A historiografia medieval portuguesa na viragem do milénio: análise bibliométrica (2000-2010) e representação taxonómica* [Electronic Version]. Retrieved 22 oct. 2014, from <https://dspace.uevora.pt/rdpc/handle/10174/11220?mode=full>

Moreiro González, José Antonio (2006). La representación y recuperación de los contenidos digitales: de los tesauros conceptuales a las folksonomías. In: Tramullas, J. (coord.). *Tendencias en Documentación Digital*. Gijón: TREA: 81-109.

Moreiro González, José Antonio (2011). *Linguagens documentárias e vocabulários semânticos para a web: elementos concetuais*. Salvador: EDUFBA. ISBN 978-85-232-0824-0.

NP 4036 (1992). Documentação. *Tesauros monolíngues: directivas para a sua construção e desenvolvimento*. Lisboa: Instituto Português da Qualidade.

Portugal. Biblioteca Nacional. Área de Indexação e Classificação (1998). *SIPORBASE: sistema de indexação em português: manual*. 3.^a ed. rev. e aum. Lisboa: Biblioteca Nacional. ISBN 972-565-154-5.

Simões, Maria da Graça (2008). *Da abstracção à complexidade formal: as relações conceptuais num tesouro*. Coimbra: Almedina. ISBN 972-98827-0-3.

Zhonghong, Wang; Chaudhry, Abdus Sattar; Khoo, Cristopher (2006). Potencial and prospects of taxonomies for content organization. *Knowledge organization: international journal devoted to concept theory, classification, indexing and knowledge representation*, vol. 33 (3), 160-170. ISSN 0943-7444.